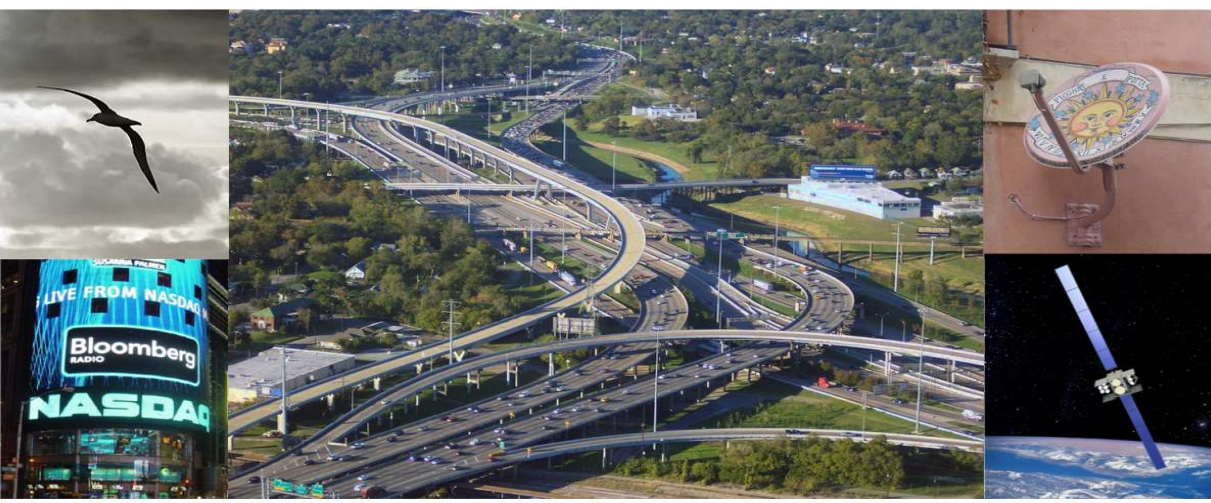
KATHOLIEKE UNIVERSITEIT
LEUVEN

Arenberg Doctoral School of Science, Engineering & Technology
Faculty of Science
Department of Physics and Astronomy

Statistical mechanics of spin models on graphs

Izaak Neri



Dissertation presented in partial fulfillment of the requirements for the degree of Doctor in Physics

June 2010

Statistical mechanics of spin models on graphs

Izaak Neri

Jury:
Prof. Dr. Jos Rogiers, president
Prof. Dr. Désiré Bollé, promotor
Prof. Dr. Andrea Pagnani
Prof. Dr. Christian Maes
Prof. Dr. Enrico Carlon
Prof. Dr. Joseph Indekeu

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor in physics

D/2010/10.705/38

May 2009

D/2010/10.705/38
ISBN 978-90-8649-345-6

Preface

The aim of fundamental research is to give a simple description of a whole set of reproducible phenomena appearing around us which allows to predict future outcomes of experiments. In theoretical physics this description is given by mathematical equations which we call laws of nature.

The laws of Newton predict the trajectory of an apple falling from a tree and the moon circling around the earth, two seemingly distinct problems. Another beautiful example of the predictive power of such fundamental theories is given by statistical mechanics. In the 19-th century Ludwig Boltzmann realized that all macroscopic objects around us are made of many small entities called molecules. The reaction of these macroscopic objects to forces, temperature gradients or particle flows are described in a very simple way by statistical laws. Seemingly distinct problems as the stability of stars and ferromagnetism are described by statistical mechanics.

The underlying structures of traffic networks, neural networks, the internet can all be described by graphs. We would like to study processes on these graphs such as traffic jams, memory recovery and routing processes. These processes can be modelled by spin models on graphs, which find their origin in the study of magnetic systems. A powerful predictive framework is given by the belief-propagation or cavity method. Seemingly different problems ranging from error-correction to the prediction of protein functions and the appearence of complexity in optimization problems are all studied using this framework. This thesis is a small contribution to the statistical mechanics of spin models on graphs.

Acknowledgements

A Ph.D. is not the work of one person. It is the accomplishment resulting from the combined effort of several people which deserve to be thanked now.

First I want to express my great gratitude to the persons I have directly worked with. Many thanks go to my supervisor Désiré for all the support he has given me throughout the years. It feels good to have a strong supervisor standing as a pillar behind what you do. I am grateful for the freedom you have given me and the trust you have shown in me. Second, I warmly thank Nikos for transferring to me the passion for finitely connected systems. I was really impressed by how much joy you could retrieve from doing calculations. At last I also want to thank Fernando for the beautiful way we have complemented each other in research.

The second group of people are less visible but have still been very important to the accomplishment of this work. Thank you Filip and Anneleen for your continuous support. Many thanks also to Bavo, Anita and Christine. It has been a luxury to work with such competent staff members.

Thanks to Rob, Myriam, Fernando and Nick for making it every day a pleasant idea to go to work. Thank you Yves to have endured me as an office mate.

I thank the official members of the jury for their contribution to the scientific defense in a critical but also a respectful way. I thank especially Andrea for coming all the way from Turin to Leuven for my Ph.D. defense.

A last word goes to all the kind people I have met in Leuven. I am astonished by how many people have shown patience, openness and friendship towards me. These are traits I really admire.

Ik zou mijn ouders en mijn broer willen bedanken omdat ik weet dat voor hun deze thesis niet zo belangrijk is. Dank u wel om mij het gevoel te geven dat mijn eigen welzijn het allerbelangrijkst is.

This thesis belongs to all of you!

Abstract

Statistical mechanics of finitely-connected systems is the study of spin models on random graphs. The methods developed to study spin models on graphs find their origins in spin-glass theory and are applicable to the description of emergent phenomena in coding theory, computer science, statistical inference in molecular biology, etc.

The cornerstone of spin-glass theory is the Sherrington-Kirkpatrick model. This is a fully-connected model of spin variables interacting through random interactions drawn from a Gaussian distribution. The Sherrington-Kirkpatrick model is universal in the sense that the phase diagram of every fully-connected model with a finite variance is the same as the phase diagram of the Sherrington-Kirkpatrick model. When the variance of the distribution of couplings is infinitely large we speak of the Lévy spin glass. We show how this fully-connected model Lévy spin glass can be decomposed in a finitely-connected backbone of strong bonds interacting with the background of weak bonds. On the basis of this decomposition we derive the phase diagram of the Lévy spin glass.

We show how to extend the tools of statistical mechanics of finitely-connected systems to non-equilibrium models where the steady state is not known a priori. Our aim is the development of an efficient and accurate message-passing algorithm for the study of non-equilibrium models, analogous to the belief-propagation algorithm for equilibrium models. We construct such an algorithm through the analysis of some simple models: models of binary spins evolving through Glauber dynamics on partially asymmetric graphs which are local tree like. Non-equilibrium systems evolving on a complex network are e.g. neurons in biological organisms, the spreading of epidemics and traffic networks. We discuss the influence of the graph structure on the retrieval properties of a neural network.

Spin models on graphs can be used to study the performance of low-density parity-check codes. Low-density parity-check codes are efficient codes reaching the Shannon limit for reliable communication over memoryless symmetric channels. An important question is whether the low-density parity-check codes can also reach the Shannon limit in asymmetric channels are finite-state Markov channels. Finite

state Markov-channels or channels with memory in the noise. These channels are more realistic models for optical and wireless communication. For asymmetric channels the code performance depends on the codeword sent while for channels with memory the performance depends on the state during the use of the channel. The decoding of low-density parity-check code for channels with memory is realized through a message passing algorithm on a small-world hypergraph.

Beknopte samenvatting

De statistische mechanica van eindig geconnecteerde systemen bestaat uit de studie van spinmodellen op willekeurige grafen. De methoden die ontwikkeld zijn voor de studie van spinmodellen op grafen vinden hun oorsprong in de theorie van spinglazen, i.e. magneten met willekeurige interacties. Deze methoden zijn toepasbaar in de beschrijving van emergente fenomenen in coderingstheorie, informatica, moleculaire biologie, e.a.

De hoeksteen van de theorie van spinglazen is het Sherrington-Kirkpatrick model. Dit is een volledig geconnecteerd model voor spinvariabelen die interageren door middel van willekeurige interacties getrokken uit een Gaussische verdeling. Het Sherrington-Kirkpatrick model is universeel in de zin dat het fasediagram van elk volledig geconnecteerde model met interacties getrokken uit een kansverdeling met een eindige variantie hetzelfde is als het fasediagram van het Sherrington-Kirkpatrick model. Wanneer de variantie van de verdeling van de interacties oneindig is spreekt men van een Lévy spinglas. We tonen hoe een Lévy spinglas kan opgedeeld worden in een eindig geconnecteerde component van sterke interacties en een achtergrondruis komende van de zwakke interacties. Via deze decompositie bepalen we het fasediagram van Lévy spinglazen.

We breiden de caviteitsmethode voor de statistische mechanica van eindig geconnecteerde systemen in evenwicht uit tot niet-evenwichtsmodellen waarbij de stationaire toestand a priori niet gekend is. We hebben een "message-passing" algoritme ontwikkeld voor niet-evenwichtssystemen analoog aan het "belief-propagation" algoritme voor systemen in evenwicht via de studie van simpele modellen: Ising modellen die evolueren via een Glauber dynamica op een partieel asymmetrische graaf die een lokale boomstructuur bezit. Niet-evenwichtssystemen zijn bijvoorbeeld neuronen in biologische organismen, het verkeer en de evolutie van een epidemie. We bespreken de invloed van de graafstructuur op de performantie van herkenning van patronen door het neurale netwerk.

Wij hebben de performantie van "low-density parity-check" codes bepaald door ze te modeleren met spinmodellen op grafen. "Low-density parity-check" codes zijn

efficiënte codes die de Shannon limiet bereiken voor betrouwbare communicatie over symmetrische kanalen. Een belangerijke vraag is of men de Shannon limiet kan bereiken met “low-density parity-check” codes voor asymmetrische kanalen of voor kanalen met geheugeneffecten. Deze kanalen modeleren op een meer realistische manier, respectievelijk, optische communicatie en draadloze communicatie. In asymmetrische kanalen is de performantie van de code afhankelijk van het verzonden codewoord. De kanalen met geheugen effecten in de ruis worden gemodelleerd door “finite-state Markov” kanalen. De performantie van de code hangt af van de toestanden van het kanaal gedurende het versturen van de codewoorden. De kanalen met geheugeneffecten hebben we gemodelleerd met behulp van een small-world hypergraaf.

Nomenclature

Notations:

- $\{x_i\}_{i=1..N} = \{x_1, x_2, \cdots, x_N\}$
- $x_U = \{x_i | i \in U\}$
- $[1, N]_{\mathbb{N}} = [1, N] \subset \mathbb{N}$
- $\Omega^N = \Omega \times \Omega \cdots \times \Omega$, with Ω a set
- $\mathcal{N}\left(a, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)$
- $\sum_{i \neq j}^{N} A_{ij} = \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} - \sum_{i=1}^{N} A_{ii}$
- $\mathbb{Z}_2 = \{0, 1\}$
- $\mathbb{Z}_2^N = \{0, 1\} \otimes \{0, 1\} \cdots \{0, 1\}$
- $\int dx f(x) = \int_{-\infty}^{+\infty} dx f(x)$
- $p_{\mathrm{p}}(k; c) = \frac{\exp(-c)c^k}{k!}$
- $1..n = 1, ..., n$

Acronyms:

- 1RSB: one step replica symmetry breaking
- ABP: asymmetric belief propagation
- AWGN: additive white Gaussian noise
- BAC: binary asymmetric channel
- BEC: binary erasure channel

- BP: belief propagation
- BSC: binary symmetric channel
- CH: chaotic
- CLT: central limit theorem
- ECC: error-correcting code
- F: ferromagnetic
- FC: fully-connected
- FiC: finitely-connected
- i.i.d.r.v. : identically and independent-distributed random variables
- M: mixed
- MPA: message passing algorithm
- P: paramagnetic
- REM: random energy model
- RCM: random codeword model
- RS: replica-symmetric
- RSB: replica-symmetry breaking
- SG: spin glass
- SK: Sherrington-Kirkpatrick

Cover picture

A compilation of pictures taken from (top to bottom and left to right):
http://en.wikipedia.org/wiki/File:Albatross_shape.png
http://en.wikipedia.org/wiki/File:NASDAQ.JPG
http://en.wikipedia.org/wiki/File:45intoI-10_2.jpg
http://en.wikipedia.org/wiki/File:Taormina_-_Parabola_satellitare-meridiana_-_Gennaio_2006_Foto_di_Giovanni_Dall
http://en.wikipedia.org/wiki/File:Wideband_Global_SATCOM_Satellite.jpg

The files have the following licences:

Creative Commons Attribution ShareAlike 3.0 license
Creative Commons Attribution ShareAlike 3.0 license
Creative Commons Attribution ShareAlike 3.0 license
Creative Commons Attribution-Share Alike 2.5 Italy license
I, the copyright holder of this work, hereby release it into the public domain. This applies worldwide.

Il faut que la peinture serve à autre chose qu'à la peinture
René Magritte

Contents

Chapter 1

Introduction

The study of the collective behavior of interacting molecules, magnetic moments or electrons is not possible from a deterministic description of all the consisting units. However, a statistical theory provides a simple microscopic description of such systems. This theory founded by Boltzmann, Gibbs and others in the late 1800s is called statistical mechanics and describes emergent phenomena, such as magnetism and superconductivity. Emergence is the appearance of macroscopic order or the sudden change of the macroscopic properties of a system from the interplay of simple microscopic interactions.

It is a challenge to apply the ideas of statistical mechanics to problems which lie beyond the "traditional physics" problems. Emergent phenomena appear in computer science, information theory, biology and economics, e.g. the appearance of complexity in optimization problems, the failure of reliable communication through a noisy channel, protein folding and traffic congestion. When mapping the interactions of the consisting units of these problems on a graph we find complex interaction structures in comparison to the simple Bravais lattice we obtain after mapping the interactions of magnetic moments in a crystalline solid state on a graph. Neuronal networks, protein-protein interaction networks, networks of decoding variables, transportation networks are all examples of such networks. Analogous to the study of magnetism through spin models on a Bravais lattice one maps these problems on spin models on a graph.

The methods developed in the last 20 years to study spin models on graphs find their origins in theories that describe the phase of a magnetic material containing a highly non-trivial kind of order, a spin glass. The magnetic moments of a spin glass freeze in different, apparently random directions. To get more insight in this intricate phase, physicists developed "simple" models which can be solved exactly: the mean-field models. The mathematical and computational tools that were

developed to solve these mean-field models allow physicists to solve spin models on graphs.

In this context we have studied three problems: error-correcting codes, dynamics of spin models on graphs and Lévy spin glasses. These are three different problems but they are all situated in the context of statistical mechanics of Ising models on graphs.

Before explaining these problems we present in this introduction some historical context and mention some recent developments in the statistical mechanics of Ising models on graphs. We have made a small selection from the large literature available on finite connectivity (FiC) systems.

1.1 Spin glasses

Spin glasses are a state of condensed matter found in alloys that consist of a noble metal host diluted with magnetic transition metal impurities. Examples are materials like CuMn, AuFe. At low temperatures one discovers a cusp in the susceptibility but no apparent order, i.e. the magnetization is zero. The susceptibility is the derivative of the magnetization with respect to an external field. The basic physics of these alloys are grasped by a simple model with Hamiltonian

$$H\left(\{\sigma_i\}_{i=1..N}\right) = -\sum_{(i,j)\in\Lambda^d} J_{ij}\sigma_i\sigma_j \,, \tag{1.1}$$

with Λ^d the set of bonds on a d-dimensional Bravais lattice. The spin variables $\sigma_i \in \{-1, 1\}$ present the two states of the magnetic moments of the transition metal impurities. The coupling strengths J_{ij} are random variables fixed during the whole experiment. We call such variables *quenched variables*.

When the interactions J_{ij} fulfill the following two criteria:

1. the interactions are drawn from a random distribution,
2. this random distribution has a finite support on both negative and positive values,

a new non-trivial phase can appear at low temperatures T. This phase is called the spin-glass phase. When one removes one of the two above mentioned criteria, much of the important thermodynamic and dynamical characteristics of the spin-glass phase disappear.

In the spin-glass phase the material has no long range order: $\lim_{r\to\infty}\langle\sigma_i\sigma_{i+r}\rangle_{\rm th} \to 0$, with $\langle\cdot\rangle_{\rm th}$ the thermal average with respect to a certain experimental setup.

The order in this phase can be measured through the Edward-Anderson order parameter $q_{\mathrm{EA}} = N^{-1}\sum_{i=1}^{N}\langle\langle\sigma_i\rangle_{\mathrm{th}}^2\rangle_J$, which reflects a freezing of the spins in random directions. The outer brackets denote an average over the interaction variables while the inner brackets denote the thermal average. There exist many different configurations into which the spins can freeze. This slows down the dynamics of spin glasses.

To find the properties of this spin-glass phase, Sherrington and Kirkpatrick introduced in their seminal paper [161] a long-ranged version of (1.1):

$$H\left(\{\sigma_i\}_{i=1..N}\right) = -\sum_{i\neq j}^{N}\frac{J_{ij}}{2}\sigma_i\sigma_j\ . \tag{1.2}$$

The interactions are i.i.d.r.v. drawn from the normal distribution $\mathcal{N}\left(\frac{J_0}{N},\frac{J_1}{\sqrt{N}}\right)$. This problem is very difficult: finding the configurations $\{\sigma_i\}_{i=1..N}$ that minimize the Hamiltonian (1.2) is an NP-complete problem [9]. These configurations are important at low temperatures. The number of local minima in the free energy landscape scales exponentially in the system size. One speaks of non-trivial ergodicity breaking in contrast with the ergodicity breaking present in ferromagnetic systems where the free energy landscape contains a finite number of minima. The SK-model has been solved exactly by Parisi in [137, 138, 106] using the replica method. This exact solution reveals an intricate phase with an order parameter $P(q)$, the distribution of overlap parameters $q \equiv N^{-1}\sum_{i=1}^{N}\langle\sigma_i\rangle_\alpha\langle\sigma_i\rangle_\beta$ between two minima α,β of the free energy landscape.

Studying spin glasses leads to insights in the phenomenon of non-trivial ergodicity breaking which is not solely present in physics. The many valley structure of SG's is relevant in phenomena in biology, computer science and information theory. For an extensive explanation on spin glasses and its theories we refer to the book [54].

1.2 Statistical mechanics of Ising models on graphs

Viana and Bray [169] introduced a more realistic mean-field model for a spin glass (than the SK-model):

$$H\left(\{\sigma_i\}_{i=1..N}\right) = -\sum_{(i,j)\in E}J_{ij}\sigma_i\sigma_j\ , \tag{1.3}$$

with E the set of edges of a Poissonian graph drawn from the ensemble $\mathcal{G}_{\mathrm{p}}(N,c)$, see appendix B.2.2, and J_{ij} i.i.d.r.v. drawn from some distribution P_{I}. This model is more realistic since every spin interacts with a finite neighborhood of spins ∂i.

When the mean connectivity c of the Poissonian graph becomes infinite, one finds the SK-model.

The Ising model on a Poissonian graph is a specific example of a whole family of spin models which can be solved exactly. These mean-field models are defined on graphs with a minimal typical loop length of the order $\mathcal{O}(\log(N))$ in the system size N. We call these models also finite connectivity (FiC) systems. At high temperatures one can solve these models exactly in the thermodynamic limit. The order parameter of these mean-field models is infinite dimensional containing all the overlap parameters q_n,

$$q_n = N^{-1} \sum_{i=1}^{N} \langle \sigma_k \rangle_{\alpha_1} \langle \sigma_i \rangle_{\alpha_2} \cdots \langle \sigma_i \rangle_{\alpha_n}, \tag{1.4}$$

between $n \in \mathbb{N}$ replicas of the system. At low temperatures a scheme, analogous to the Parisi-scheme used in the SK model, has been suggested [113, 104]. The exact solution of FiC systems at low T is in general not known because it is more difficult to deal with an infinite dimensional order parameter. With the cavity method one can derive an algorithm to determine the marginals and the thermodynamic quantities on a given graph instance [104, 105].

1.3 An open path to interdisciplinary research

Many problems in interdisciplinary research can be cast into a statistical mechanics problem on a graph. Indeed, identification of protein functions in protein-protein interaction networks, evaluation of decoding algorithms in error-correcting codes, solving optimization problems, finding the eigenvalues of random matrices, interactions of humans in social networks are all examples of problems defined on a complex network. The underlying units are correlated through this network. In these problems we do not know the precise value of these units but we know the total probability distribution of the configurations of the units. These problems are very similar to the study of the spin-glass model on a random graph. To be a little bit more concrete we give some examples.

1.3.1 Applications to biology

Inference algorithms on graphs are used on networks in biology.

Gene regulatory networks

The human body contains many cells with different functions determined by the presence of different proteins. However, the "source code", i.e. the DNA, is the same in every cell. To know the source code is not sufficient to understand how the human body works. One has to understand how the "compiler" works, which is presented by the gene regulatory network.

The chemical composition of proteins is encoded in the DNA. Gene expression is the synthesis of proteins via DNA (central dogma of molecular biology [33, 34]) which is regulated by the presence of other proteins. This complex interplay between genes, proteins and other components is to a large extend modelled by a gene regulatory network. With the gene regulatory network of a cell one can start to study cell differentiation and the influence of drugs on this process. We give a small overview of possible applications:

- For some simple organisms, such as E. coli [160] and baker's yeast [69, 109] the gene regulatory network is known. In [31] the gene regulatory mechanism is modelled by a boolean network. One can study the fixed points of these networks and the genes that control these fixed points (the computational core).
- To determine experimentally the underlying gene regulatory network of an organism is in general difficult. Gene regulatory networks are only known for simple organisms. One would like to infer the underlying gene regulatory network from results obtained from DNA microarrays [20].
- Given a data set of expressed genes coming from DNA microarrays of tumor cells and normal cells, one would like to infer a sparse signature of genes from which one can identify cancer cells. This sparse signature can be used to reduce the noise of experiments and data analysis. In [136] a linear classifier, i.e. a perceptron [75], has been inferred from data coming from microarrays using a message passing algorithm.

Protein-protein interaction network

Protein-protein interaction networks contain the information of experimentally tested interactions among two proteins in graph like form [52, 53, 51]. Databases like "MIPS" [2] and "Gene Ontology" [1] contain the different functions of these proteins. Using a cost function that maximizes the number of interacting proteins with similar functions, the unidentified functions of proteins in the protein-protein interaction network have been identified [90]. The method has been tested on the protein-protein interaction network of Saccaromices Cerevisiae [52].

Figure 1.1: A schematic overview of the processes in error-correcting codes. The message $\boldsymbol{s}_0$ is encoded in the word $\boldsymbol{s}$ sent through the protocol $\boldsymbol{G}$. The decoder determines an estimate $\hat{\boldsymbol{s}}$ of the sent word $\boldsymbol{s}$ given the received word $\boldsymbol{r}$ and the encoding process $\boldsymbol{G}$.

1.3.2 Applications to information theory

An example of a scientific interplay between information theory and statistical mechanics are error-correcting codes. Error-correcting codes are constructed to make reliable communication possible between two persons connected by a noisy channel. The receiver has to infer the sent message $\boldsymbol{s}$ from a corrupted version $\boldsymbol{r}$, given the encoding procedure G. This is possible through a decoding algorithm that uses the predefined correlations between the variables of the message introduced in the encoding process. The encoding procedure is a map G from the message $\boldsymbol{s}_0$ to the word $\boldsymbol{s}$ sent. The word sent contains many redundant bits to counter the effect of noise. A possible decoding approach is to estimate the bits through $s_i = \operatorname{sign} \sum_s s P_i\left(s|\boldsymbol{r}, G\right)$, with P_i the i-th marginal of the posterior probability distribution. A schematic overview is given in figure 1.1. In general it is very hard to find the marginals since one has to compute a summation containing 2^{N-1} terms.

The rate is defined as the ratio of information bits to message bits. When the rate is low, one is sending a small amount of information through the channel but the message becomes less prone to errors. Therefore, one wants to maximize the rate under the condition that reliable communication is possible. Shannon proved that this maximum is given by the capacity [159], a quantity which depends on the channel and not on the precise encoding or decoding procedure. In coding theory one tries to find practical, i.e. computationally efficient, encoding and decoding procedures that reach the Shannon limit. This problem has been solved using a decoding procedure consisting of calculating marginals on a FiC. In figure 1.2 we show the graphical model used for decoding algorithms in LDPC codes [61, 94] and Turbo codes [14, 120], two of the best performant codes.

Besides their theoretical interests LDPC and Turbo codes are also used in practical applications, for instance in deep-space exploration where the strength of the signal over long distances is really important. Turbo codes have been used in the Pathfinder mission [23] and Turbo codes and LDPC codes have been introduced

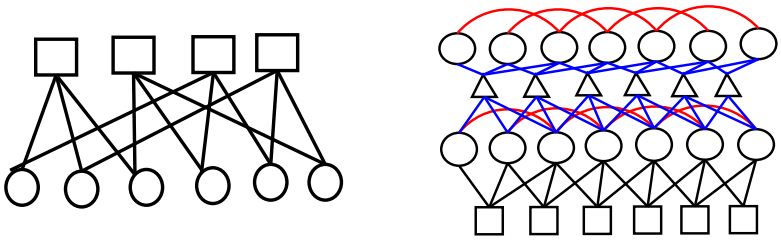


Figure 1.2: The bipartite Tanner graph used in LDPC codes (left) and the graph for Turbo codes (right)

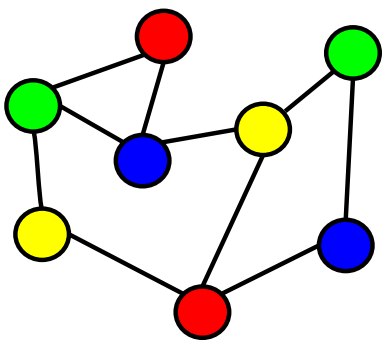


Figure 1.3: A proper cover of 4 colors on the given graph.

in protocols for digital video broadcasting [23].

1.3.3 Applications to computer science

Consider the following problem. Given a graph, is it possible to color the vertices with q different colors such that any two adjacent nodes have different colors? In figure 1.3 we present an example of a proper coloring of a graph. The graph coloring problem is NP-complete. This means that the problem is very difficult: there is no known algorithm that can solve the problem with a computational complexity scaling polynomially in the number of vertices N. The q-coloring is one of many known NP-complete problems: K-SAT and the travelling-salesman problem to mention a few others. The classification in hard (NP problems) and easy problems (P problems) is based on a worst-case analysis. For a random graph drawn from, for instance the Erdős-Rényi ensemble, the problem is in general easy. To be more precise, as a function of the mean graph connectivity c, for $N \to \infty$ the problem undergoes at a certain point c_s a phase transition from a colorable region where graphs are colorable with probability one to an uncolorable region where graphs

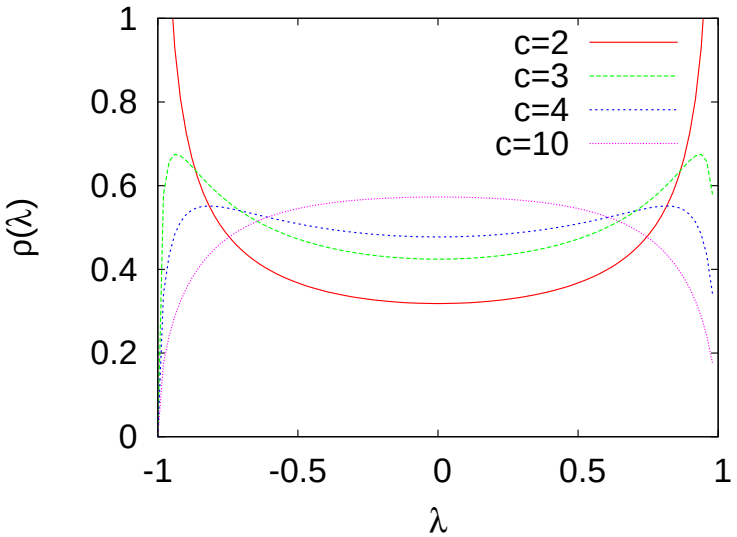


Figure 1.4: The distribution of eigenvalues ρ of the coordination matrices of Bethe lattices for a given fixed connectivity c. Spectra of sparse matrices have been found by mapping the problem on a spin model on a graph. For $c \to \infty$ the spectra converge to the Wigner semi-circle law.

are uncolorable with probability one. Around the threshold c_s the computational complexity of the problem increases [26, 154, 155]. With the cavity method [178] it is possible to study the geometrical structure of the set of solutions and try to understand the reason for the computational hardness of this optimization problem. It was conjectured in [178] that the increase in computational complexity around the phase transition is associated with the rigidity threshold c_r. In the rigid region $c_r < c < c_s$ the entropic dominant solutions contain a non-zero fraction of frozen variables.

1.3.4 Applications to mathematics

The statistics of eigenvalues of $N \times N$ random matrices, for $N \to \infty$, is an example of a distribution of many correlated random variables. Random matrices are of interest in financial applications [141], electronic models [27], nuclear models [100], etc. In the case that the matrix elements are i.i.d.r.v. drawn from a Gaussian, one retrieves the Wigner semi-circle law. For sparse matrices, i.e. every row and column contains a finite number of non-zero elements, the eigenvalue spectra can be calculated using a $\mathcal{O}(N)$ algorithm based on the cavity

method. This algorithm outperforms standard diagonalization methods which have a computational complexity $\mathcal{O}(N^3)$ [142].

The method was first used in [29] for random Lévy matrices and later applied to sparse random matrices [150]. The central quantity in these calculations is the *resolvent*:

$$G(z) = (z - A)^{-1}. \tag{1.5}$$

We define $A^{(i)}$ as the submatrix of A where we removed the i-th column and i-th row and $G^{(i)}$ as its resolvent. From the inverse of a partitioned matrix [76], one finds the following expression:

$$\frac{1}{G_{ii}(z)} = z - A_{ii} - \sum_{l,k(\neq i)} A_{ik} G^{(i)}_{kl}(z) A_{li}. \tag{1.6}$$

The off-diagonal terms in the summation have no significant contribution [29]. When one removes j on the left and right hand side and assumes $G^{(i,j)}_{kl}(z) = G^{(i)}_{kl}(z)$, one gets,

$$\frac{1}{G^{(j)}_{ii}(z)} = z - A_{ii} - \sum_{k(\neq i,j)} |A_{ik}|^2 G^{(i)}_{kk}(z). \tag{1.7}$$

These are precisely the cavity equations for random matrices as derived in [150]. From the definition of the resolvent one finds:

$$\operatorname{Tr} G(z) = \int d\lambda \frac{\rho(\lambda)}{z - \lambda}, \tag{1.8}$$

and thus $\rho(\lambda) = \pi^{-1} \lim_{\epsilon \to 0^+} \lim_{N \to \infty} \operatorname{Im} \operatorname{Tr} G(z)$, with $\epsilon = \operatorname{Im} z$. The eigenvalue spectrum of the coördination matrices of Bethe lattices is given by

$$\rho(\lambda) \sim \frac{\sqrt{1 - \lambda^2}}{\frac{c^2}{4(c-1)} - \lambda^2}. \tag{1.9}$$

When $c \to \infty$ one recovers the Wigner semi-circle law $\rho(\lambda) = \frac{2}{\pi}\sqrt{1 - \lambda^2}$. These spectra are plotted in figure 1.4.

1.4 Problems discussed in this thesis

In this thesis we discuss three different problems related to statistical mechanics of FiC systems.

- *Lévy spin glasses*: we study the generalization of the SK-model to coupling distributions that do not fulfill the central limit theorem. This fully-connected model turns out to be a hybrid between a fully-connected (FC) and a FiC model.
- *Dynamics of Ising models on graphs*: we discuss the use of the cavity method to the parallel dynamics of Ising models on graphs. We construct a message passing algorithm to determine the marginals of the stationary distribution of a non-equilibrium model.
- *Low-density parity-check codes*: we make a study of LDPC codes on asymmetric channels and channels with memory. This study extends on previous calculations made on symmetric channels.

1.5 Outline of the thesis

In the introduction we have situated the research of statistical mechanics of Ising models on random graphs. In the second chapter we describe the main technical calculations on a straightforward model familiar to all physicists: the Ising model. The first two chapters make the thesis readable and interesting to a general physics audience. Experts in the field can skip these two chapters.

The following three chapters can all be read independently and contain the new results obtained. These chapters contain their own introduction, conclusion and future prospects. The technical details are kept to a minimum as they are analogous (but more complicated) to the ones explained in the second chapter. The focus is put on the results and their implications for these problems. In the third chapter we determine the phase diagram of Lévy spin glasses. In the fourth chapter we study the parallel dynamics of Ising models on graphs. In the fifth chapter we explain the implementation of LDPC codes to asymmetric channels and channels with memory.

Chapter 2

Techniques for statistical mechanics of spin models on graphs

This chapter is an introduction to the statistical mechanics of FiC systems. Two methods to study spin models on graphs, the replica method and the cavity method, are explained. Using the replica method one determines the thermodynamic behavior of typical graphs drawn from some ensemble of graphs. The cavity method on the other hand studies the thermodynamic behavior of a spin model on a given graph instance. We discuss how spontaneous symmetry breaking and spin-glass phases are described by these methods. At the end of the chapter, we present the phase diagrams of the SK model, the Viana-Bray model, p-spin models and of the random energy model and elaborate on the nature of the SG phase in these models. For a more thorough discussion we refer the book of Mézard and Montanari [102].

2.1 Spin models on graphs

Central problems

We consider a given mass probability function $P(\{\sigma_i\}_{i=1..N})$, with $\sigma_i \in \Omega$. The canonical ensemble in equilibrium statistical mechanics is determined by the distribution

$$P\left(\{\sigma_i\}_{i=1..N}|\alpha\right) \sim \exp\left[-\beta H\left(\{\sigma_i\}_{i=1..N}|\alpha\right)\right], \tag{2.1}$$

with the cost function $H\left(\{\sigma_i\}_{i=1..N}|\alpha\right)$. For physical systems, like magnets or fluids, H is the energy of the system and the quantity $\beta = T^{-1}$ is the inverse temperature. We will also call less correctly H the Hamiltonian. With α we denote all the external parameters on which the Hamiltonian can depend, such as the external field. We refer to the distribution (2.1) as the Gibbs distribution. We want to solve the following questions in the thermodynamic limit when the system size N becomes infinitely large:

- We want to calculate the different moments of the distribution P. The first moment is called the magnetization m:

$$m \equiv \lim_{N\to\infty} \sum_{\{\sigma_i\}_{i=1..N}} \left(\frac{\sum_{i=1}^{N} \sigma_i}{N} \right) P\left(\{\sigma\}_{i=1..N}|\alpha\right). \tag{2.2}$$

  The n-th moment of P is called the n-point correlation function.

- We want to calculate the entropy S ($s = S/N$) which is related to the "volume-size" of the typical set in Ω^N with respect to P:

$$s \equiv \lim_{N\to\infty} -\frac{1}{N} \sum_{\{\sigma_i\}_{i=1..N}} P\left(\{\sigma\}_{i=1..N}|\alpha\right) \log\left(P\left(\{\sigma\}_{i=1..N}|\alpha\right)\right). \tag{2.3}$$

  The Boltzmann constant k_B, appearing in the definition of the entropy in statistical physics [68], is set to one.

- The average energy E ($e = E/N$) is defined through:

$$e \equiv \lim_{N\to\infty} -\frac{1}{N} \sum_{\{\sigma_i\}_{i=1..N}} P\left(\{\sigma\}_{i=1..N}|\alpha\right) H\left(\{\sigma_i\}_{i=1..N}|\alpha\right). \tag{2.4}$$

- In many practical problems one would like to know the single site marginals P_i:

$$P_i(\sigma_i|\alpha) = \lim_{N\to\infty} \sum_{\sigma_{[1,N]_{\mathbb{N}}\setminus i}} P\left(\{\sigma\}_{j=1..N}|\alpha\right). \tag{2.5}$$

The quantities m, s and e are the thermodynamic values of the magnetization, the entropy and the energy. The summations in (2.2-2.5) contain $|\Omega|^N$ terms and are in general very hard to perform. We are concerned in finding efficient algorithms to perform these summations. The magnetization and the entropy are macroscopic

and extensive quantities. One calculates them by taking the derivative of the free energy F, ($f = F/N$), with respect to its coördinates:

$$F(\beta, \alpha) \equiv -\beta^{-1} \log \left(\sum_{\{\sigma_i\}_{i=1..N}} \exp\left[-\beta H\left(\{\sigma_i\}_{i=1..N} | \alpha\right)\right] \right). \tag{2.6}$$

In general, phase transitions appear because of two "competing" effects in the thermodynamic limit as a function of the temperature: minimization of the energy at low temperatures and maximization of the entropy at high temperatures. Indeed, the free energy is the Legendre-Fenchel transform of the entropy:

$$f(T) = \min_e \left(e - Ts(e)\right). \tag{2.7}$$

In this thesis we will only study cases where $\Omega = \{-1, 1\}$. We speak of Ising models and we call the variables σ_i Ising variables. A generalization to q-spins or continuous variables is mostly straightforward, although the resulting equations are more complicated.

Graphs and ensembles of graphs

We consider a specific class of mass probability functions P where the correlations between the different variables σ_i can be presented through a graph $G = (E, V)$:

$$P\left(\{\sigma_i\}_{i=1..N} | G, \boldsymbol{J}\right) = \prod_{a \in E} \psi\left(\{\sigma_i\}_{i \in a} | J_a\right), \tag{2.8}$$

with E the set of edges of the given graph G and $\boldsymbol{J}$ the matrix of weights defined on the set E. In appendix B one can find the definitions and basic ideas from graph theory that we will use in this thesis.

In practice the graph G is obtained from experiments. It can be the gene regulatory network of a yeast cell [69, 109] or the transportation network of global cargo ship movements [81] or some other network structure. The probability distribution P is determined by, respectively, the biological processes or the transportation processes one wants to describe with the network

The specific graph structure has a big influence on the macroscopic behavior of these models. The study of Ising models on a given graph instance G is in general a very hard problem, see equations (2.2-2.5). We study Ising models on graphs drawn from a certain graph ensemble $\mathcal{G}$. We will consider graphs drawn from the Erdős-Rényi ensemble, the Poissonian ensemble, etc. (see appendix [**?**]). The properties of typical graphs, drawn from the probability distribution $P_{\mathcal{G}}(G)$ of the

ensemble $\mathcal{G}$, are well understood and therefore a perfect ground for theoretical research. In theoretical models we draw the interactions $\boldsymbol{J}$ from a distribution $P_I(\boldsymbol{J})$. We leave the subindices $\mathcal{G}$ and I in P away when there is no confusion possible.

We have two types of random variables: the quenched variables $\boldsymbol{J}$ and G and the annealed variables $\{\sigma_i\}_{i=1..N}$. We are interested in the macroscopic properties of the probability distribution $P(\{\sigma_i\}_{i=1..N}|G,\boldsymbol{J})$ for a typical $(G,\boldsymbol{J})$ and not in the macroscopic properties of the joint distribution $P(\{\sigma_i\}_{i=1..N},G,\boldsymbol{J})$. From a dynamical point of view the spins can vary in time but the interactions can not.

Self-averaging

As the graph structure and its interactions are random the free energy f is also a random variable. We limit ourselves to the much easier calculation of the average free energy $\overline{f} = \sum_{G,\boldsymbol{J}} P(G)P(\boldsymbol{J})f(G,\boldsymbol{J})$.

In general we call a random quantity X self averaging when:

$$\lim_{N\to\infty} P_{X,N}\left(\left|X-\overline{X}\right|>\epsilon\right)\to 0,\ \forall\epsilon>0, \tag{2.9}$$

with ϵ an arbitrary small positive number. The use of random quantities that concentrate for $N\to\infty$ around their mean value is essential in our calculations.

The calculation of the average magnetization, the average entropy and other extensive quantities, see subsection 2.1, is meaningful thanks to the self-averaging properties of the corresponding random variables. An "experiment" consists of the sampling of a very small set of configurations $\{\sigma_i\}_{i=1..N}$ from the distribution P. To make useful predictions, through the calculation of averages, on the measured values of the observables, we are interested in observables that for $N\to\infty$ contain a certain value with probability one. We call this value the typical value of the observable. The self-averaging of the magnetization in statistical physics follows from the Einstein fluctuation theory and the extensivity of the entropy [24]. The importance of the self-averaging effect is also present in large deviation theory descriptions of statistical mechanics [168].

Completely analogous, one demands the self-averaging of the free energy f. Indeed, in general the number of samples drawn from $P(G,\boldsymbol{J})$ is much smaller then the sample space. For some particular models, like the SK-model and the Viana-Bray model, self-averaging has been proven [55, 71]. The self-averaging of the free energy $f(G,\boldsymbol{J})$ is not proven in general for mean-field models. In this thesis we assume the self-averaging conjecture to be valid during all the calculations. Therefore, it is important to compare theoretical predictions with simulations.

Ising model on an ensemble with a given degree distribution

We derive the thermodynamic quantities for an Ising model on a given graph instance G drawn from the ensemble $\mathcal{G}_{\rm deg}$ with the replica method and the cavity method. This is a simple model whose solution illustrates the basic ingredients used in more complicated calculations of more realistic models. The probability distribution in this case is:

$$P\left(\{\sigma_i\}_{i=1..N}|\boldsymbol{C},\boldsymbol{J}\right) \sim \exp\left[-\beta H\left(\{\sigma_i\}_{i=1..N}|\boldsymbol{C},\boldsymbol{J}\right)\right], \tag{2.10}$$

with the Hamiltonian:

$$H(\{\sigma_i\}_{i=1..N}|\boldsymbol{C},\boldsymbol{J}) \equiv -\sum_{(i,j)} c_{ij}J_{ij}\sigma_i\sigma_j. \tag{2.11}$$

The quantities $\boldsymbol{C} = \{c_{ij}\}_{i,j=1..N}$, $c_{ij} \in \{0,1\}$ (see appendix C), and $\boldsymbol{J} = \{J_{ij}\}_{i,j=1..N}$ denote, respectively, the connectivity matrix and the matrix of the interactions. The couplings $\boldsymbol{J}$ are i.i.d.r.v. drawn from a distribution P_I, i.e.:

$$P_I(\boldsymbol{J}) = \prod_{i<j} P_I(J_{ij}). \tag{2.12}$$

The graph G is drawn from the ensemble $\mathcal{G}_{\rm deg}$, determined by a degree distribution $p_{\rm d}$, see appendix B.2.3. The probability of the graphs in this ensemble is given by

$$P_{\rm deg}(G|N,p_{\rm d}(k)) = \prod_{i=1}^{N} p_{\rm d}(k_i) P_{\rm deg}(\boldsymbol{C}|\{k_i\}_{i=1..N}), \tag{2.13}$$

and

$$P_{\rm deg}(\boldsymbol{C}|\{k_i\}_{i=1..N}) \sim \left(\prod_{i<j}\frac{c}{N}\delta(c_{ij};1) + \left(1-\frac{c}{N}\right)\delta(c_{ij};0)\right) \times \prod_{i=1}^{N}\delta\left(k_i;\sum_j c_{ij}\right). \tag{2.14}$$

with $\boldsymbol{C}$ the connectivity matrix of the graph. The graphs drawn from $\mathcal{G}_{\rm deg}$ have typically few loops and the degrees of neighbouring vertices are uncorrelated. These two properties, allow us to solve spin models on these graphs for $N \to \infty$.

From these solutions we can develop useful algorithms. Sometimes it is convenient to add external fields $\boldsymbol{\theta}$ to the Hamiltonian (2.11)

$$H(\{\sigma_i\}_{i=1..N}|\boldsymbol{C},\boldsymbol{J},\boldsymbol{\theta}) \equiv -\sum_{(i,j)} c_{ij}J_{ij}\sigma_i\sigma_j - \sum_{i=1}^{N}\theta_i\sigma_i. \tag{2.15}$$

2.2 Replica method

The central idea of the replica method is to perform the average of the free energy over the quenched variables before the summation over the spin values. With the replica method one calculates the disorder-averaged free energy $\overline{f}$ using a certain "predefined/straightforward" scheme. It is therefore no coincidence that most mean-field models [7, 38, 161, 169] have been solved first with this method. The most beautiful result obtained by the replica method is with no doubt the exact solution of the SG phase of the SK model [106, 137, 138, 164].

2.2.1 The free energy

Most of the important statistical information is contained in the free energy f:

$$\begin{aligned} f(\boldsymbol{C},\boldsymbol{J}) &\equiv -\lim_{N\to\infty}\frac{\log(Z)}{N\beta} \\ &= -\lim_{N\to\infty}\frac{\log\left(\sum_{\{\sigma_i\}_{i=1..N}}\exp\left[-\beta H(\{\sigma\}_{i=1..N}|\boldsymbol{C},\boldsymbol{J})\right]\right)}{N\beta}. \end{aligned} \tag{2.16}$$

Equation (2.16) defines the partition function Z. Assuming self-averaging of the free energy f, we calculate the disorder-averaged value $\overline{f}$:

$$\overline{f} = -(N\beta)^{-1}\overline{\log\left(\sum_{\{\sigma\}_{i=1..N}}\exp\left[-\beta H\left(\{\sigma_i\}_{i=1..N}|\boldsymbol{C},\boldsymbol{J}\right)\right]\right)}, \tag{2.17}$$

with $\overline{f}$ the average value of the free energy with respect to the quenched variables $\boldsymbol{k}$, $\boldsymbol{C}$ and $\boldsymbol{J}$:

$$\overline{f} = \sum_{\boldsymbol{k},\boldsymbol{C},\boldsymbol{J}} p_{\mathrm{d}}(\boldsymbol{k})P_{\mathrm{deg}}(\boldsymbol{C}|\{k_i\}_{i=1..N})P_I(\boldsymbol{J})f(\boldsymbol{C},\boldsymbol{J}). \tag{2.18}$$

We leave out the subindex in $p_{\rm d}$ when no confusion is possible.

We take the average of the free energy over the quenched variables with the replica trick [107]:

$$\overline{\log Z} = \lim_{n\to 0} \frac{\overline{Z^n - 1}}{n} = \lim_{n\to 0} \log \frac{\overline{Z^n}}{n}. \tag{2.19}$$

We apply this trick to (2.17), for n a natural number, by replicating the summation n-times. The σ_i variables are replicated to σ_i^α variables with $\alpha = 1..n$. As a general rule we use Latin indices for sites and Greek indices for replicas. The former run from $1..N$ while the latter from $1..n$. The sum over the $\{c_{ij}\}_{i=1..N, j=1..N}$-variables can be taken in the expression of Z^n using the expression (B.12) for $P_{\rm deg}$:

$$\begin{aligned} -\beta \overline{f} = \lim_{N\to\infty} \lim_{n\to 0} \log \frac{1}{N\mathcal{M}} \sum_{\{\sigma_i^\alpha\}} \sum_{\{k_i\}} \prod_i p(k_i) \int \left(\prod_i \frac{d\omega_i}{2\pi} \right) e^{i \sum_i \omega_i k_i} \\ \prod_{i<j} \exp\left[\frac{c}{N} \left(e^{-i\omega_i} e^{-i\omega_j} \langle e^{\beta J \sum_\alpha \sigma_i^\alpha \sigma_j^\alpha} \rangle_J - 1 \right) \right], \end{aligned} \tag{2.20}$$

where we used

$$\delta_{k_i, \sum_j c_{ij}} = \int_0^{2\pi} \frac{d\omega_i}{2\pi} \exp\left[i\omega_i (k_i - \sum_j c_{ij}) \right]. \tag{2.21}$$

The quantity $\mathcal{M}$ in equation (2.20) is the normalization constant of $P_{\rm deg}(\boldsymbol{C} | \{k_i\}_{i=1..N})$, whose value is given by (B.13). In the argument of the exponential of expression (2.20) we insert the identity

$$1 = \sum_{\boldsymbol{\sigma}} \delta_{\boldsymbol{\sigma}, \boldsymbol{\sigma}_i}, \tag{2.22}$$

with $\boldsymbol{\sigma} = (\sigma^1, \sigma^2, \cdots, \sigma^n)$ to get

$$
\begin{aligned}
-\beta \overline{f} &= \lim_{N\to\infty} \lim_{n\to 0} \log \frac{1}{\mathcal{M}} e^{-N\frac{c}{2}} \sum_{\{k_i\}} \prod_i p(k_i) \int \left(\prod_i \frac{d\omega_i}{2\pi} \right) e^{i \sum_i \omega_i k_i} \\
&\quad \times \sum_{\{\sigma_i^\alpha\}} \exp \left[\frac{Nc}{2} \sum_{\boldsymbol{\sigma},\boldsymbol{\tau}} \left(\frac{\sum_i e^{-i\omega_i} \delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}_i}}{N} \right) \left(\frac{\sum_j e^{-i\omega_j} \delta_{\boldsymbol{\tau},\boldsymbol{\sigma}_j}}{N} \right) \right. \\
&\quad \left. \times \langle e^{-\beta J \sum_\alpha \sigma^\alpha \tau^\alpha} \rangle_J + \mathcal{O}(1) \right]. \qquad (2.23)
\end{aligned}
$$

The order parameters $P(\boldsymbol{\sigma})$ are defined through

$$
P(\boldsymbol{\sigma}) \equiv \frac{\sum_i \delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}_i} e^{-i\omega_i}}{N}, \qquad (2.24)
$$

and are added explicitly to (2.23) through the following identity:

$$
1 = \int \prod_{\boldsymbol{\sigma}} \left[\frac{dP(\boldsymbol{\sigma}) d\hat{P}(\boldsymbol{\sigma})}{2\pi N} \right] \exp \left[iN\hat{P}(\boldsymbol{\sigma}) \left(P(\boldsymbol{\sigma}) - \frac{\sum_i \delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}_i} e^{-\omega_i}}{N} \right) \right]. \qquad (2.25)
$$

The free energy becomes

$$
\begin{aligned}
-\beta \overline{f} &= \lim_{N\to\infty} \lim_{n\to 0} \frac{1}{N} \log \left\{ \exp \left[N \left(\frac{c}{2} - c \log c + \sum_{k\geq 0} p(k) \log k! \right) \right] \right. \\
&\quad \left. \times \int \prod_{\boldsymbol{\sigma}} \left[\frac{dP(\boldsymbol{\sigma}) d\hat{P}(\boldsymbol{\sigma})}{2\pi N} \right] \exp \left[N\Psi \left(P(\boldsymbol{\sigma}), \hat{P}(\boldsymbol{\sigma}) \right) \right] \right\}, \qquad (2.26)
\end{aligned}
$$

with

$$
\begin{aligned}
\Psi \left(P(\boldsymbol{\sigma}), \hat{P}(\boldsymbol{\sigma}) \right) &= \frac{c}{2} \sum_{\boldsymbol{\sigma},\boldsymbol{\tau}} P(\boldsymbol{\sigma}) P(\boldsymbol{\tau}) \langle e^{-\beta J \sum_\alpha \sigma^\alpha \tau^\alpha} \rangle_J \\
&\quad + i \sum_{\boldsymbol{\sigma}} \hat{P}(\boldsymbol{\sigma}) P(\boldsymbol{\sigma}) + \log \left(\sum_k p(k) \sum_{\boldsymbol{\sigma}} \int \frac{d\omega}{2\pi} e^{i\omega k} e^{-i\hat{P}(\boldsymbol{\sigma}) e^{-i\omega}} \right). \qquad (2.27)
\end{aligned}
$$

When we interchange the two limits, i.e. we perform the limit $\lim_{N\to\infty}$ before $\lim_{n\to 0}$, the expression of the average free energy $\overline{f}$ becomes a steepest-descent

integral, see appendix A.2, leading to

$$-\beta\overline{f} = \lim_{n\to 0}\frac{1}{n}\left(\frac{c}{2}-c\log c+\sum_{k\geq 0}p(k)\log k!+\frac{c}{2}\sum_{\boldsymbol{\sigma},\boldsymbol{\tau}}P^*(\boldsymbol{\sigma})P^*(\boldsymbol{\tau})\langle e^{\beta J\sum_\alpha \sigma^\alpha\tau^\alpha}\rangle_J\right.$$
$$\left.+i\sum_{\boldsymbol{\sigma}}\hat{P}^*(\boldsymbol{\sigma})P^*(\boldsymbol{\sigma})+\log\left(\sum_k p(k)\sum_{\boldsymbol{\sigma}}\int\frac{d\omega}{2\pi}e^{\omega k}e^{-i\hat{P}^*(\boldsymbol{\sigma})e^{-i\omega}}\right)\right), \quad (2.28)$$

with $\hat{P}^*(\boldsymbol{\sigma})$ and $P^*(\boldsymbol{\sigma})$ the values of $P(\boldsymbol{\sigma})$ and $\hat{P}(\boldsymbol{\sigma})$ at the saddle point. We eliminate the conjugate variable $\hat{P}^*(\boldsymbol{\sigma})$ in $\overline{f}$:

$$-\beta\overline{f} = \lim_{n\to 0}\frac{1}{n}\left(\frac{c}{2}-c\log c+\sum_{k\geq 0}p(k)\log k!-\frac{c}{2}\sum_{\boldsymbol{\sigma},\boldsymbol{\tau}}P(\boldsymbol{\sigma})P(\boldsymbol{\tau})\langle e^{\beta J\sum_\alpha \sigma^\alpha\tau^\alpha}\rangle_J\right.$$
$$\left.+\log\left(\sum_k p(k)\sum_{\boldsymbol{\sigma}}\int\frac{d\omega}{2\pi}e^{i\omega k}\exp\left[ce^{-i\omega}\sum_{\boldsymbol{\tau}}P(\boldsymbol{\tau})\langle e^{\beta J\sum_\alpha \sigma^\alpha\tau^\alpha}\rangle_J\right]\right)\right), \quad (2.29)$$

with $P(\boldsymbol{\sigma})$ the solution of

$$P(\boldsymbol{\sigma}) = \sum_k p(k)\frac{\int\frac{d\omega}{2\pi}e^{i\omega(k-1)}\exp\left[ce^{-i\omega}\sum_{\boldsymbol{\tau}}P(\boldsymbol{\tau})\langle e^{\beta\sum_\alpha J\sigma^\alpha\tau^\alpha}\rangle_J\right]}{\sum_k p(k)\sum_{\boldsymbol{\sigma}}\int\frac{d\omega}{2\pi}e^{i\omega k}\exp\left[ce^{-i\omega}\sum_{\boldsymbol{\tau}}P(\boldsymbol{\tau})\langle e^{\beta\sum_\alpha J\sigma^\alpha\tau^\alpha}\rangle_J\right]}. \quad (2.30)$$

We have left out the superscript in P^*. After performing the integrations over the ω-variables the free energy becomes:

$$-\beta\overline{f} = \lim_{n\to 0}\frac{1}{n}\left(\frac{c}{2}-\frac{c}{2}\sum_{\boldsymbol{\sigma},\boldsymbol{\tau}}P(\boldsymbol{\sigma})P(\boldsymbol{\tau})\langle e^{\beta J\sum_\alpha \sigma^\alpha\tau^\alpha}\rangle_J\right.$$
$$\left.+\log\left[\sum_k p(k)\sum_{\boldsymbol{\sigma}}\left(\sum_{\boldsymbol{\tau}}P(\boldsymbol{\tau})\langle e^{-\beta J\sum_\alpha \sigma^\alpha\tau^\alpha}\rangle_J\right)^k\right]\right), \quad (2.31)$$

with P a solution of

$$P(\boldsymbol{\sigma}) = \sum_{k=0}^{\infty} \frac{p(k)k}{c} \frac{\left(\sum_{\boldsymbol{\tau}} P(\boldsymbol{\tau}) \langle e^{\beta J \sum_{\alpha} \sigma^{\alpha}\tau^{\alpha}} \rangle_J\right)^{k-1}}{\sum_k p(k) \sum_{\boldsymbol{\sigma}} \left(\sum_{\boldsymbol{\tau}} P(\boldsymbol{\tau}) \langle e^{\beta J \sum_{\alpha} \sigma^{\alpha}\tau^{\alpha}} \rangle_J\right)^{k}}. \tag{2.32}$$

The paramagnetic (P) solution $P(\boldsymbol{\sigma}) = 2^{-n}$ always fulfills equations (2.32). To find other solutions, corresponding with spin-glass (SG) and ferromagnetic (F) phases we have to take the $\lim_{n\to 0}$ using some physical intuition. This amounts to making an ansatz for the order parameters $P(\boldsymbol{\sigma})$ and $\hat{P}(\boldsymbol{\sigma})$.

2.2.2 Replica symmetry

Density evolution equations

To take the limit $n \to 0$ one needs to make an analytic continuation of the result (2.31) and (2.32) to real values of n. The choice of the analytic continuation depends on some physical intuition. The simplest approach is using the RS ansatz [83]. We introduce a distribution of fields $W_c(h)$, just like has been done in [103, 113]:

$$P_{\mathrm{RS}}(\boldsymbol{\sigma}) = \int dh W_c(h) \prod_{\alpha} \frac{\exp(\beta h \sigma^{\alpha})}{2\cosh(\beta h)}. \tag{2.33}$$

We notice that $P_{\mathrm{RS}}(\boldsymbol{\sigma}) = P_{\mathrm{RS}}(\pi\boldsymbol{\sigma})$ and $\hat{P}_{\mathrm{RS}}(\boldsymbol{\sigma}) = \hat{P}_{\mathrm{RS}}(\pi\boldsymbol{\sigma})$, with $\pi \in \mathcal{S}_N$ a permutation. The function $\hat{P}_{\mathrm{RS}}$ is defined through the saddle point equations derived from (2.26). The distribution of $W_c(h)$ is equal to the distribution of cavity fields, hence we have introduced the subscript c, see the next section 2.3. Substitution of (2.33) in the saddle point equations (2.32) leads to the result

$$\int dh W_c(h) \prod_{\alpha} \frac{\exp(\beta h \sigma^{\alpha})}{2\cosh(\beta h)}$$

$$= \sum_k \frac{p(k)k}{c} \frac{\langle \prod_{\alpha} \prod_{r=1}^{k-1} \sum_{\tau} \exp(\beta J_r \tau \sigma^{\alpha}) \frac{\exp(\beta h_r \tau)}{2\cosh(\beta h_r)} \rangle_{h_{1..k-1},;J_{1..k-1}}}{\sum_k p(k) \langle \left(\sum_{\sigma} \prod_{r=1}^{k} \sum_{\tau} \exp(\beta J_r \tau \sigma) \frac{\exp(\beta h_r \tau)}{2\cosh(\beta h_r)}\right)^{n} \rangle_{h_{1..k-1};J_{1..k-1}}}, \tag{2.34}$$

where we used the notation

$$\langle f(h_{1..k}, J_{1..k})\rangle_{h_{1..k};J_{1..k}} = \int \left(\prod_{r=1}^{k} dJ_r P(J_r)\right)\left(\prod_{r=1}^{k} dh_r W_c(h_r)\right) f(h_{1..k}, J_{1..k}).$$

After we have taken the limit $n \to 0$ we get the self-consistent equation for W_c

$$W_c(h) = \sum_{k=0}^{\infty} \frac{p(k)k}{c} \int \prod_{r=1}^{k-1} dJ_r P(J_r) \int \prod_{r=1}^{k-1} dh_r W_c(h_r)$$
$$\times\delta\left[h - \beta^{-1}\sum_{r=1}^{k-1} \operatorname{atanh}\left(\tanh\left(\beta h_r\right)\tanh\left(\beta J_r\right)\right)\right]. \tag{2.35}$$

The equations (2.35) are called the *density evolution equations*. The order parameters like the magnetization m are found as

$$m = \int dh W(h)\tanh(\beta h), \tag{2.36}$$

with W the distribution of real fields fulfilling

$$W(h) = \sum_{k=0}^{\infty} p(k) \int \prod_{r=1}^{k} dJ_r P(J_r) \int \prod_{r=1}^{k} dh_r W_c(h_r)$$
$$\times\delta\left(h - U\left(\{h_r\}_{r=1..k}, \{J_r\}_{r=1..k}\right)\right), \tag{2.37}$$

where we have introduced the function

$$U\left(\{h_r\}_{r=1..k-1}, \{J_r\}_{r=1..k-1}\right) \equiv \beta^{-1}\sum_{r=1}^{k-1} \operatorname{atanh}\left(\tanh\left(\beta h_r\right)\tanh\left(\beta J_r\right)\right). \tag{2.38}$$

For a Poissonian graph we have $W_c(h) = W(h)$ since $\frac{k}{c}p_{\mathrm{p}}(k;c) = p_{\mathrm{p}}(k-1;c)$.

Free energy

One obtains the RS expression $f_{\rm RS}$ for the free energy $\overline{f}$ after substitution of the RS ansatz (2.33) in (2.29):

$$-\beta f_{\rm RS} = \lim_{n\to 0}\frac{1}{n}\left(\frac{c}{2} - \frac{c}{2}\left\langle\left(\sum_{\sigma,\tau}\exp\left(\beta J\sigma\tau\right)\frac{\exp\left(\beta h\sigma+\beta g\tau\right)}{4\cosh\left(\beta h\right)\cosh\left(\beta g\right)}\right)^n\right\rangle_{h,g;J}\right.$$
$$\left. + \log\left[\sum_k p(k)\left\langle\left(\sum_\sigma\prod_{r=1}^k\left(\sum_\tau\frac{\exp\left(\beta h_r\tau\right)}{2\cosh\left(\beta h_r\right)}\exp\left(\beta J_r\sigma\tau\right)\right)\right)^n\right\rangle_{h_{1..k};J_{1..k}}\right]\right) \tag{2.39}$$

We take the limit $n\to 0$ using the two identities:

$$\lim_{n\to 0}\left\langle\frac{\left(f(x_1,x_2,\cdots,x_l)\right)^n-1}{n}\right\rangle_{x_1,\cdots,x_l} = \left\langle\log f(x_1,x_2,\cdots,x_l)\right\rangle_{x_1,\cdots,x_l},$$
$$\lim_{n\to 0} n^{-1}\log\left\langle f^n(x_1,x_2,\cdots,x_l)\right\rangle_{x_1,\cdots,x_l} = \left\langle\log f(x_1,x_2,\cdots,x_l)\right\rangle_{x_1,\cdots,x_l},$$

with $\langle\cdot\rangle_{x_1,\cdots,x_l}=1$ and f some function. The RS expression $f_{\rm RS}$ equation (2.39) becomes

$$f_{\rm RS} = \langle\Delta F_{\rm site}\rangle_{h_{1..k};J_{1..k}} - \frac{c}{2}\langle\Delta F_{\rm link}\rangle_{h,g;J}, \tag{2.40}$$

with

$$-\beta\Delta F_{\rm link} = \log\left(\sum_{\sigma,\tau}\exp\left(\beta J\sigma\tau\right)\frac{\exp\left(\beta h\sigma+\beta g\tau\right)}{4\cosh\left(\beta h\right)\cosh\left(\beta g\right)}\right), \tag{2.41}$$

$$-\beta\Delta F_{\rm site} = \sum_k p(k)\log\left[\sum_\sigma\prod_{r=1}^k\left(\sum_\tau\frac{\exp\left(\beta h_r\tau\right)}{2\cosh\left(\beta h_r\right)}\exp\left(\beta J_r\sigma\tau\right)\right)\right]. \tag{2.42}$$

We end this subsection with some remarks:

- The original problem of a summation over the set of variables $\{\sigma_i\}_{i=1..N}$ (2.2-2.5), is reduced to finding the solution to equation (2.35). This is a recursive distributional equation, i.e. an equation of the type $W = T(W)$, with W a probability distribution on $\mathbb{R}$, and $T : \mathcal{W} \to \mathcal{W}$ a map from $\mathcal{W}$, the set of all probability distribution on $\mathbb{R}$, to itself. One can compare these equations with the fixed point equations of the Curie-Weiss model $m = \tanh(\beta J m)$, with $W_c(h) = \delta(h; \beta J)$. Studies on the existence, uniqueness, stability and the domain of attraction of the fixed points to these equations can be found in [6], and references therein.

- A simple example of a recursive distributional equations is:

$$W(h) = \int dh_1 dh_2 W(h_1) W(h_2) \delta\left(h - \frac{h_1 + h_2}{\sqrt{2}}\right) . \tag{2.43}$$

  The equation (2.43) contains infinitely many fixed points given by $\mathcal{N}\left(\mu, \sigma^2\right)$. The domain of attraction of these fixed point equations is given by $\mu = \int dh W_0(h) h$ and $\sigma^2 = \int dh W_0(h) h^2 - \mu^2$.

2.2.3 The different phases

We consider three different phases depending on the values of the magnetization m and the SG order parameter q:

$$m = \int dh W(h) \tanh(\beta h), \tag{2.44}$$

$$q = \int dh W(h) \tanh^2(\beta h). \tag{2.45}$$

These phases are:

- the P phase: the equations (2.35) admit the solution $W_c(h) = \delta(h)$ for all temperatures. We call this the P phase.

- the F phase: a F solution is a solution with $m > 0$ and $q > 0$. In the left part of figure 2.1 we present the F solution for the Ising model on a Poissonian graph with mean connectivity $c = 3$ and $P_I(J) = \delta(J - 1/c)$. At $T = 0$ the distribution $W_c(h)$ is a sum of delta functions corresponding to the different possible values of the fields, i.e. a/c with $a \in \mathbb{N}$.

- the SG phase: a SG solution is a solution with $q > 0$ and $m = 0$. In the right part of figure 2.1 we show a SG solution for the Ising model on a Poissonian graph with mean connectivity $c = 3$. The couplings are unbiased

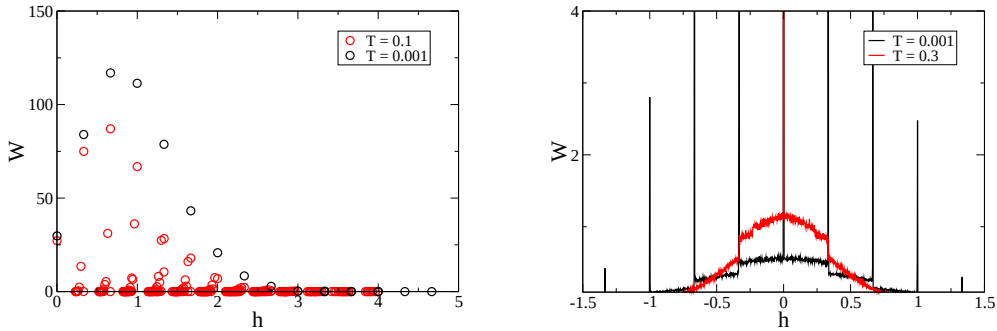


Figure 2.1: The order parameter W for a Poissonian graph with mean connectivity $c = 3$. Left: the couplings have all the same value $J = 1/c$. Right: the couplings have equal probability to have the value $1/c$ and $-1/c$.

and bimodal, i.e. they have equal probability to have the values $-1/c$ and $1/c$. The distribution $W_c(= W)$ in the SG phase contains two parts: a discrete and a continuous part. For $T \to 0$ the continuous part does not disappear. The reason for this non intuitive result is that the RS ansatz becomes incorrect at low temperatures in the SG phase.

Phase transitions: bifurcation lines

At high temperatures the equations (2.35) only admit the P solution. At low temperatures the SG or F solutions appear. One finds the critical temperatures $\beta_{\rm F}$ and $\beta_{\rm SG}$, from the P solution to, respectively, the F or SG solution, using an expansion around $W_c(h) = \delta(h)$:

$$\mathrm{P-F}: \quad 1 = \left(\sum_{k=0}^{\infty} \frac{k}{c}(k-1)p(k)\right) \langle \tanh\left(\beta_{\rm F} J\right)\rangle_J, \tag{2.46}$$

$$\mathrm{P-SG}: \quad 1 = \left(\sum_{k=0}^{\infty} \frac{k}{c}(k-1)p(k)\right) \langle \tanh^2\left(\beta_{\rm SG} J\right)\rangle_J. \tag{2.47}$$

We make a distinction between dynamic (or spinodal) transitions and thermodynamic (or static) transitions. The former are identified by the stability of solutions to the density evolution equations (2.35). The thermodynamic transitions are determine by the values of the averages over the Gibbs measure (2.10). These values are determined by the solution to (2.35) with the lowest or the highest free energy (2.40).

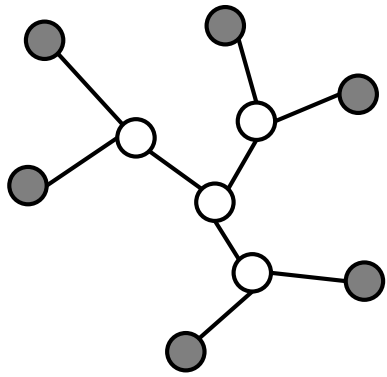


Figure 2.2: A Cayley tree of size $N = 10$ is presented. The internal vertices are empty while the vertices in the leaves are filled. Internal vertices have a degree $k = 3$. Vertices in the leaves of the graph have a degree $k = 1$. The degree distribution of Cayley trees for $N \to \infty$ is given by $p_{\rm d}(k) = \frac{z-1}{z}\delta(k;1) + \frac{1}{z}\delta(k;z+1)$, with $z+1$ the degree of the internal vertices. Remark that there are non-trivial degree-degree correlations in the Cayley tree.

When we set $p(k) = \delta(k - C)$ we recover the critical inverse temperature $\tanh(\beta_{\rm F} J) = (C-1)^{-1}$ for the Ising model on the Bethe lattice [11]. We do not recover the thermodynamic behavior of the Ising model on a Cayley tree. The Ising model on the Cayley tree has no spontaneous symmetry breaking [45, 123], in contrast with the finite critical temperature given by (2.46). However, the Cayley tree contains non-trivial degree-degree correlations in the leaves of the graph. Therefore it is not a graph typically drawn from the ensemble (2.13-2.14).

Population dynamics

The F-SG transition has to be determined numerically, since no analytic solutions are known for the F solution and SG solution. One can solve (2.35) using a Monte Carlo method [104]. The unknown distribution $W_c(h)$ is presented by a population of $\mathcal{N}$ fields. This population is updated sequentially through the following algorithm which performs one sweep, with input parameters the number of fields $\mathcal{N}$, the population of $\mathcal{N}$ fields W_c, the degree distribution $p(k)$, the distribution of couplings $P_I(J)$ and the mean connectivity c:

```
popDyn(W, N, p(k), P(J), c)[
        for(int i=1..N)[
                *draw a degree k from the distribution p(k)k/c
```

```
                *draw k-1 fields from the population W of N fields
                *draw k-1 couplings from
                 the distribution  of couplings P(J)
                *calculate the new field hT with
                 the function U defined in equation (2.35)
                *draw uniformly an integer i between 1 and N
                *Replace W[i] with hT
        ]
]
```

The distributions in the figures (2.1) are computed using the population dynamics algorithm for $\mathcal{N} \sim \mathcal{O}(1e+6)$.

We end this paragraph with a remark. In practice the population dynamics algorithm is performed with a truncated degree distribution $\tilde{p}_{\mathrm{d}}$

$$\tilde{p}_{\mathrm{d}}(k) \sim p_{\mathrm{d}}(k)\theta\left(K-k\right). \tag{2.48}$$

For Poissonian degree distributions the phase diagram of the truncated model converges exponentially fast to the asymptotical phase diagram $K \to \infty$ since $p_{\mathrm{p}}(k) \sim \mathcal{O}(\exp\left(-k\log(k)\right))$. If we take a power law distribution $p_{\mathrm{d}}(k) \sim k^{-\lambda}$, which is quite common in the literature [42, 82, 91], the error of truncation is not small anymore. In chapter 4 we will show how to control the error when discussing the dynamics of a neural network on a scale-free graph.

2.2.4 Replica symmetry breaking

The RS ansatz does not always lead to exact results: we say that the RS ansatz becomes unstable. Replica symmetry breaking (RSB) is very important since the RS is broken in the SG phase. The RSB corresponds with a complex free energy landscape which makes the SG phases essentially different from the F phase. RSB can be local or non-local and a myriad of equivalent methods exist to analyze the RSB-instability. An extensive and accurate explanation of RSB in FiC system can be found in the following Ph. D. theses [149, 177].

Local symmetry breaking

- **Eigenvalues of the Hessian:**
  RSB follows from the instability of the RS ansatz (2.33) to RSB fluctuations. The steepest descent integral (2.26) is only valid at a point in the $\left\{P(\boldsymbol{\sigma}), \hat{P}(\boldsymbol{\sigma})\right\}$-phase space where the eigenvalues of the Hessian are all

positive, see subsection A.2 in appendix A. When the Hessian $\mathcal{H}$ at the RS point $(P_{\rm RS}(\boldsymbol{\sigma}), \hat{P}_{\rm RS}(\boldsymbol{\sigma}))$,

$$\mathcal{H} = \begin{pmatrix} \left.\frac{\partial^2 \Psi}{\partial P(\boldsymbol{\sigma}) \partial P(\boldsymbol{\tau})}\right|_{P=P_{\rm RS}, \hat{P}=\hat{P}_{\rm RS}} & \left.\frac{\partial^2 \Psi}{\partial P(\boldsymbol{\sigma}) \partial \hat{P}(\boldsymbol{\tau})}\right|_{P=P_{\rm RS}, \hat{P}=\hat{P}_{\rm RS}} \\ \left.\frac{\partial^2 \Psi}{\partial \hat{P}(\boldsymbol{\sigma}) \partial P(\boldsymbol{\tau})}\right|_{P=P_{\rm RS}, \hat{P}=\hat{P}_{\rm RS}} & \left.\frac{\partial^2 \Psi}{\partial \hat{P}(\boldsymbol{\sigma}) \partial \hat{P}(\boldsymbol{\tau})}\right|_{P=P_{\rm RS}, \hat{P}=\hat{P}_{\rm RS}} \end{pmatrix} \quad (2.49)$$

gets a negative eigenvalue, RSB is broken. The quantity Ψ is defined in (2.27) and $P_{\rm RS}$ is the RS order parameter given by (2.33) and (2.35). This instability argument was first used by Almeida and Thouless for the SK model [36]. We are not aware of such an analysis for FiC.

- **Two-replica method:**
We consider, according to [89], two uncoupled replicas defined on the same graph with the same interactions, i.e. a model with a Hamiltonian given by

$$H\left(\{\sigma_i\}_{i=1..N}, \{\tau_i\}_{i=1..N}\right) = -\sum_{(i,j)\in E} J_{ij}\left(\sigma_i\sigma_j + \tau_i\tau_j\right). \quad (2.50)$$

The order-parameter equations, within the RS assumption, are then

$$\begin{aligned} W_c\left(h^{(1)}, h^{(2)}\right) = \sum_{k=0}^{\infty} \frac{p(k)k}{c} \int \prod_{r=1}^{k-1} dJ_r P(J_r) \int \prod_{r=1}^{k-1} dh_r^{(1)} dh_r^{(2)} W_c(h_r^{(1)}, h_r^{(2)}) \\ \times\delta\left(h^{(1)} - \beta^{-1} \sum_{r=1}^{k-1} \operatorname{atanh}\left(\tanh\left(\beta h_r^{(1)}\right)\tanh\left(\beta J_r\right)\right)\right) \\ \times\delta\left(h^{(2)} - \beta^{-1} \sum_{r=1}^{k-1} \operatorname{atanh}\left(\tanh\left(\beta h_r^{(2)}\right)\tanh\left(\beta J_r\right)\right)\right). \end{aligned} \quad (2.51)$$

The RS solution is stable when the diagonal solution $W_c(h^{(1)}, h^{(2)}) = W_c(h^{(1)})\delta(h^{(1)} - h^{(2)})$ is stable under iteration through the self-consistent equation (2.51). The two-replica method is easy to use as it gives us a general way to calculate RSB-instability within a RS calculation. When one performs a stability analysis around the point $W_c(h^{(1)}, h^{(2)}) = \delta(h^{(1)})\delta(h^{(2)})$, one finds that the SG solution is always unstable.

- **The SG susceptibility:**
The SG susceptibility is defined through

$$\chi_2 = N^{-1} \sum_{i,j} \left(\langle\sigma_i\sigma_j\rangle - \langle\sigma_i\rangle\langle\sigma_j\rangle\right)^2. \quad (2.52)$$

Instability of the RS ansatz is signalled by the divergence of χ_2. A calculation of the SG susceptibility on FiC is computationally quite involved, see for instance [149, 177]. However, calculations of the SG susceptibility are quite popular in the literature.

For FC systems the three above mentioned methods are equivalent [107]. For FiC we are not aware of a proof of the equivalence between the different instability criteria.

1RSB ansatz

The RS ansatz can be stable to RSB but still lead to wrong results for the thermodynamic value of the magnetization, free energy, entropy and other thermodynamic quantities. An example is the p-spin model where the P-SG transition is discontinuous [37, 63, 66, 102]. After the P-SG transition the P solution coexists with the SG solution. The model has then 2 stable solutions.

One can find this discontinuous transition by generalizing the ansatz (2.33) to an ansatz without permutation symmetry. Following the Parisi ansatz for the SK model, we use the following ansatz as a first step in an iterative scheme of breaking the replica symmetry [113]

$$P_{\mathrm{1RSB}}(\boldsymbol{\sigma}) = \int \mathcal{D}\Pi \, W_c^{\mathrm{1RSB}}(\Pi) \prod_{\alpha=1}^{n/m} \left(\int dh \Pi(h) \prod_{\gamma=1}^{m} \frac{\exp(\beta h \sigma^{\alpha,\gamma})}{2\cosh(\beta h)} \right). \quad (2.53)$$

The integration over the distribution Π is a functional integration, i.e. an infinite dimensional integration over all the values $\Pi(h)$. The parameter m is called the Parisi RSB parameter. We also introduce $\mu = m\beta$. The meaning of the different parameters becomes clear in subsection 2.3 when we study the model with the cavity method. We substitute the 1RSB ansatz (2.53) into the saddle

point equations (2.32):

$$\int \mathcal{D}\Pi W_c^{\rm 1RSB}(\Pi) \prod_{\alpha=1}^{n/m} \left(\int dh \Pi(h) \prod_{\gamma=1}^{m} \frac{\exp(\beta h \sigma^{\alpha,\gamma})}{2\cosh(\beta h)} \right)$$
$$\sim \sum_k \frac{p(k)k}{c} \int \left(\prod_{r=1}^{k-1} \mathcal{D}\Pi^{(r)} W_c^{\rm 1RSB}(W_r) \right) \int \left(\prod_{r=1}^{k-1} dJ_r P(J_r) \right)$$
$$\times \prod_{\alpha=1}^{n/m} \left(\int \left(\prod_{r=1}^{k-1} dh_r \Pi_r(h_r) \right) \prod_{\gamma=1}^{m} \prod_{r=1}^{k-1} \frac{\exp(\beta h_r \tau^{\alpha,\gamma})}{2\cosh(\beta h_r)} \exp(\beta J_r \sigma^{\alpha,\gamma} \tau^{\alpha,\gamma}) \right) \tag{2.54}$$

We take the limit $n \to 0$ to get as final result

$$\begin{aligned} W_c^{\rm 1RSB}(\Pi) &= \sum_k \frac{p(k)k}{c} \int \left(\prod_{r=1}^{k-1} \mathcal{D}\Pi_r W_c^{\rm 1RSB}(\Pi_r) \right) \int \left(\prod_{r=1}^{k-1} dJ_r P(J_r) \right) \\ &\times \delta_F \Bigg[W(h) - 1/\mathcal{C} \int \prod_{r=1}^{k-1} dh^r \Pi_r(h^r) \exp(-\mu \Delta F_{\rm iter}) \\ &\qquad \delta\Big(h - \beta^{-1} \sum_{r=1}^{k-1} {\rm atanh}(\tanh(\beta h_r)\tanh(\beta J_r)) \Big) \Bigg] \end{aligned} \tag{2.55}$$

with $\Delta F_{\rm iter}$ given by

$$\Delta F_{\rm iter} = -\frac{1}{\beta} \log \left(\sum_\tau \prod_{r=1}^{k-1} \sum_\sigma \exp(\beta J_r \sigma \tau) \frac{\exp(\beta h_r \sigma)}{2\cosh(\beta h_r)} \right), \tag{2.56}$$

and $\mathcal{C}$ a normalization constant. Substitution of the ansatz (2.53) in (2.29) gives the 1RSB value ϕ_{1RSB} for the disorder-averaged free energy $\overline{f}$

$$\phi_{\rm 1RSB} = \sum_k p(k) \int \prod_{r=1}^{k} \mathcal{D}\Pi_c^{(r)} W_c^{\rm 1RSB}\left(\Pi_c^{(r)}\right) \int \prod_{r=1}^{k} dJ_r P(J_r)\, \Delta\Phi_{\rm site}$$
$$-\frac{c}{2} \int \mathcal{D}\Pi_c^{(1)} \mathcal{D}\Pi_c^{(2)} W_c^{\rm 1RSB}\left(\Pi_c^{(1)}\right) W_c^{\rm 1RSB}\left(\Pi_c^{(2)}\right) \Delta\Phi_{\rm link}, \tag{2.57}$$

with

$$-\mu\Delta\Phi_{\text{site}} = \log\left(\int\prod_{r=1}^{k} dh_r \Pi_c^{(r)}(h_r)\exp\left(-\mu\Delta F_{\text{site}}\right)\right), \tag{2.58}$$

$$-\mu\Delta\Phi_{\text{link}} = \log\left(\int dh_1 dh_2 \Pi^{(1)}(h_1)\Pi^{(2)}(h_2)\exp\left(-\mu\Delta F_{\text{link}}\right)\right), \tag{2.59}$$

and ΔF_{site} and ΔF_{link} as defined in (2.41) and (2.42). The model contains symmetry breaking when there is a non-trivial solution (i.e. a non P solution) for a value $\mu > 0$. We make a distinction between a continuous transition and a discontinuous transition depending on the corresponding behavior of μ (and the distribution W^{1RSB}) around the RS solution at the transition.

One finds back the RS result when one takes:
$W_c^{\text{1RSB}}(\Pi) = \int dh W_c(h)\delta_F\left(\Pi(g) - \delta(g-h)\right)$. The RS distribution W_c is found from W_c^{1RSB} through $W_c(x) = \int d\Pi W_c^{\text{1RSB}}(\Pi)\,\Pi(x)$.

2.3 Cavity method

In the cavity method one performs the summation over the spins $\{\sigma_i\}_{i=1..N}$ before the average over the quenched variables. One can take the summation over the spins by introducing cavities in the graph G, see figure 2.3. The cavity graph $G^{(i)}$ is the subgraph of G where we removed the i-th vertex and all of its edges incident to i.

The results of the replica method are recovered by the cavity method [107, 104]. Still, a derivation of the RS and 1RSB phase diagram of FiC systems with the cavity method remains interesting:

- With the cavity method one can determine macroscopic quantities on a given graph instance which is not possible with the replica method.
- The different order parameters (W_c, W_c^{1RSB}, μ) have a clear interperation in the cavity method.
- The different assumptions made in deriving the RS phase diagram become transparent in the cavity method which makes it possible to connect the cavity method with rigorous calculations [57].

In this section we solve the Ising model on the graph ensemble $\mathcal{G}_{\text{deg}}$ with the cavity method.

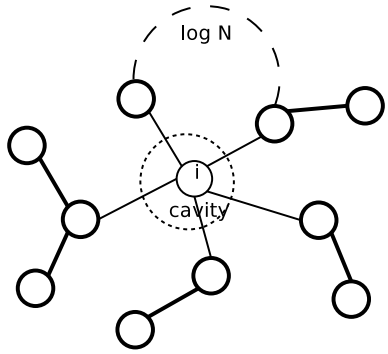


Figure 2.3: An example of the introduction of a cavity in a graph. For graphs drawn from the ensemble $\mathcal{G}_{\rm deg}$, the minimal length of loops is typically of the order $\mathcal{O}(\log N)$, with N the size of the graph. After the removal of the i-th vertex, the minimal distance between its neighbours is of the order $\mathcal{O}(\log(N))$.

2.3.1 One-pure-phase cavity method

The cavity approach is applied to the Gibbs distribution $P\left(\{\sigma_i\}_{i=1..N}\right)$ equation (2.10) with corresponding free energy F (2.6). The derivation we give below is sometimes referred to as the RS cavity method. We prefer to call it the one-pure-phase cavity method.

Graph operations

Before going into the analysis we introduce some notations for the marginal distributions of a Gibbs distribution on a graph G. The marginal probabilities $P_{\mathcal{S}}$ and $P_{\mathcal{S}}^{(\mathcal{T})}$ with $\mathcal{S}$ a subset of $[1,N]_{\mathbb{N}}$ and $\mathcal{T}$ a subset of $[1,N]_{\mathbb{N}}$ or $[1,N]_{\mathbb{N}}\times[1,N]_{\mathbb{N}}$ are

$$\begin{aligned} P_{\mathcal{S}}\left(\sigma_{\mathcal{S}}\right) &\equiv \sum_{\{\sigma_n\}_{n\in V\setminus\mathcal{S}}} \frac{\exp\left[-\beta H\left(\{\sigma_j\}_{j=1..N}|\boldsymbol{C},\boldsymbol{J}\right)\right]}{\sum_{\{\sigma_j\}_{j=1..N}}\exp\left[-\beta H\left(\{\sigma_j\}_{j=1..N}|\boldsymbol{C},\boldsymbol{J}\right)\right]}, \\ P_{\mathcal{S}}^{(\mathcal{T})}\left(\sigma_{\mathcal{S}}\right) &\equiv \sum_{\{\sigma_n\}_{n\in V\setminus\mathcal{S}}} \frac{\exp\left[-\beta H^{(\mathcal{T})}\left(\{\sigma_j\}_{j=1..N}|\boldsymbol{C},\boldsymbol{J}\right)\right]}{\sum_{\{\sigma_j\}_{j=1..N}}\exp\left[-\beta H^{(\mathcal{T})}\left(\{\sigma_j\}_{j=1..N}|\boldsymbol{C},\boldsymbol{J}\right)\right]}. \end{aligned} \tag{2.60}$$

The Hamiltonian $H^{(\mathcal{T})}$ is defined on the graph obtained from G through the corresponding graph operation labeled by $\mathcal{T}$.

The cavity approach is based on applying iteratively graph operations on the graph G and the corresponding probability P. The following operations are the most important ones:

- *Site deletion*: this process is shown graphically in figure 2.4. One deletes a site from the graph, for instance site i, and all of the links connecting this site with its neighbours. The neighbours of i become *cavity spins*. The Hamiltonian is decomposed into:

$$H = H^{(i)} + \Delta H^{(i)}_{\text{site}}, \tag{2.61}$$

which for the Hamiltonian (2.11) is given by

$$\Delta H^{(i)}_{\text{site}} \quad = \quad -\sum_{k\in\partial i} J_{ik}\sigma_i\sigma_k. \tag{2.62}$$

To $H^{(i)}$ we associate the distribution $P^{(i)}$:

$$\begin{aligned} P^{(i)}\left(\sigma_{V\setminus\overline{\partial}i};\sigma_{\partial i}\right) &= \frac{\exp\left[-\beta H^{(i)}\left(\sigma_{V\setminus\overline{\partial}i};\sigma_{\partial i}\right)\right]}{\sum_{\sigma_{V\setminus i}}\exp\left[-\beta H^{(i)}\left(\sigma_{V\setminus\overline{\partial}i};\sigma_{\partial i}\right)\right]}, \\ P^{(i)}_c\left(\sigma_{V\setminus\overline{\partial}i}|\sigma_{\partial i}\right) &= \frac{\exp\left[-\beta H^{(i)}\left(\sigma_{V\setminus\overline{\partial}i};\sigma_{\partial i}\right)\right]}{\sum_{\sigma_{V\setminus\overline{\partial}i}}\exp\left[-\beta H^{(i)}\left(\sigma_{V\setminus\overline{\partial}i};\sigma_{\partial i}\right)\right]}. \end{aligned} \tag{2.63}$$

We have made in (2.63) an explicit distinction between the normal spins $\sigma_{V\setminus\overline{\partial}i}$ and the cavity spins $\sigma_{\partial i}$. We have defined the "closed" neighbourhood of a vertex $\overline{\partial}_i \equiv \partial i \cup \{i\}$. We decompose the free energy as follows:

$$-\beta F = \log\left(\sum_{\sigma_{\overline{\partial i}}}\exp\left(-\beta F^{(i)}\left(\sigma_{\partial i}\right) - \beta\Delta H^{(i)}_{\text{site}}\right)\right) \tag{2.64}$$

with

$$\begin{aligned} F^{(i)}\left(\sigma_{\partial i}\right) &\equiv -\beta^{-1}\log\left\{\sum_{\sigma_{V\setminus\overline{\partial}i}}\exp\left[-\beta H^{(i)}\left(\sigma_{V\setminus\overline{\partial}i};\sigma_{\partial i}\right)\right]\right\} \\ &= F^{(i)}_{\text{bulk}} + F^{(i)}_{\text{cavity}}\left(\sigma_{\partial i}\right). \end{aligned} \tag{2.65}$$

The free energy $F^{(i)}$ and $F^{(i)}_{\rm bulk}$ are the free energy of respectively the distribution $P^{(i)}$ and $P^{(i)}_c$ on the "cavity" graph $G^{(i)}$. The former model contains a boundary of fixed cavity spins. The bulk free energy $F^{(i)}_{\rm bulk}$ is the constant term in the free energy, i.e. the term independent of the boundary spins $\sigma_{\partial i}$. At last we define the free energy difference $\Delta F^{(i)}_{\rm site}$, the difference in the bulk free energy of $P^{(i)}$ on $G^{(i)}$ and P on G:

$$\Delta F^{(i)}_{\rm site} \equiv F - F^{(i)}_{\rm bulk}. \tag{2.66}$$

One can apply a site deletion on the cavity graph $G^{(i)}$ with probability distribution $P^{(i)}$ leading to $G^{(i),(j)}$ and $P^{(i),(j)}$. This process can be iterated.

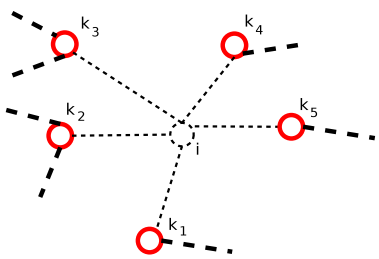


Figure 2.4: The process of site deletion. We delete the site i from the graph and all the links incident with this site. The neighbouring spins become cavity spins. We have colored them grey.

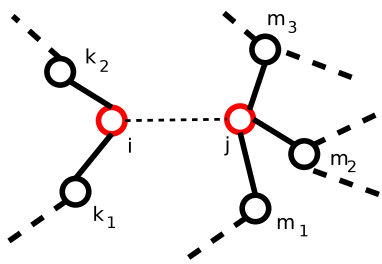


Figure 2.5: The process of link deletion. We delete a link connecting site i with site j. The i-th and the j-th site become cavity spins.

- *Link deletion*: Link deletion is the process where one deletes a link from the graph, for instance the link connecting the i-th site with the j-th site, shown in figure 2.5. The sites i and j become cavities. We define again the related changes in observables. The Hamiltonian is written as

$$H = H^{(i,j)} + \Delta H^{(i,j)}_{\rm link}, \tag{2.67}$$

with the energy difference given by

$$\Delta H^{(i,j)}_{\rm link} = -J_{ij}\sigma_i\sigma_j. \tag{2.68}$$

We associate the distribution $P^{(i,j)}$ to $H^{(i,j)}$:

$$P^{(i,j)}\left(\sigma_{V\setminus i,j};\sigma_i,\sigma_j\right)=\frac{\exp\left(-\beta H^{(i,j)}\left(\sigma_{V\setminus i,j};\sigma_i,\sigma_j\right)\right)}{\sum_{\sigma_V}\exp\left(-\beta H^{(i,j)}\left(\sigma_{V\setminus i,j};\sigma_i,\sigma_j\right)\right)}, \quad (2.69)$$

$$P_c^{(i,j)}\left(\sigma_{V\setminus i,j};\sigma_i,\sigma_j\right)=\frac{\exp\left(-\beta H^{(i,j)}\left(\sigma_{V\setminus i,j};\sigma_i,\sigma_j\right)\right)}{\sum_{\sigma_{V\setminus i,j}}\exp\left(-\beta H^{(i,j)}\left(\sigma_{V\setminus i,j};\sigma_i,\sigma_j\right)\right)}. \quad (2.70)$$

We write again

$$\begin{aligned} F^{(i,j)}(\sigma_i,\sigma_j) &= -\beta^{-1}\log\left(\sum_{\sigma_{V\setminus i,j}}\exp\left(-\beta H^{(i,j)}\left(\sigma_{V\setminus i,j};\sigma_i,\sigma_j\right)\right)\right) \\ &= F_{\text{bulk}}^{(i,j)}+F_{\text{cavity}}^{(i,j)}(\sigma_i,\sigma_j). \end{aligned} \quad (2.71)$$

The free energies $F^{(i,j)}$ and $F_{\text{bulk}}^{(i,j)}$ correspond with the probability distributions $P_c^{(i,j)}$ and $P^{(i,j)}$. The bulk free energy differences are then

$$\Delta F_{\text{link}}^{(i,j)} \equiv F-F_{\text{bulk}}^{(i,j)}. \quad (2.72)$$

- *Site iteration*: a scheme of this process is presented in figure 2.6. Site iteration is the process of site deletion on the cavity graph. An example is the deletion of the i-th site on the cavity graph $G^{(j)}$, with i and j neighbouring sites. The energy changes are given by

$$H^{(j)}=H^{(i),(j)}+\Delta H_{\text{iter}}^{(i-j)}, \quad (2.73)$$

with

$$\Delta H_{\text{iter}}^{(i-j)}=\sum_{\ell\in\partial i\setminus j}J_{\ell i}\sigma_\ell\sigma_i. \quad (2.74)$$

We have

$$P^{(i),(j)} = \frac{\exp\left[-\beta H^{(i),(j)}\left(\sigma_{V\setminus(\partial i\cup\partial j)};\sigma_{\partial i\cup\partial j}\right)\right]}{\sum_{\sigma_V}\exp\left[-\beta H^{(i),(j)}\left(\sigma_{V\setminus(\partial i\cup\partial j)};\sigma_{\partial i\cup\partial j}\right)\right]}, \quad (2.75)$$

$$P_c^{(i),(j)}\left(\sigma_{\partial i\cup\partial j}\right) = \frac{\exp\left[-\beta H^{(i),(j)}\left(\sigma_{V\setminus(\partial i\cup\partial j)};\sigma_{\partial i\cup\partial j}\right)\right]}{\sum_{\sigma_{V\setminus(\partial i\cup\partial j)}}\exp\left[-\beta H^{(i),(j)}\left(\sigma_{V\setminus(\partial i\cup\partial j)};\sigma_{\partial i\cup\partial j}\right)\right]}. \quad (2.76)$$

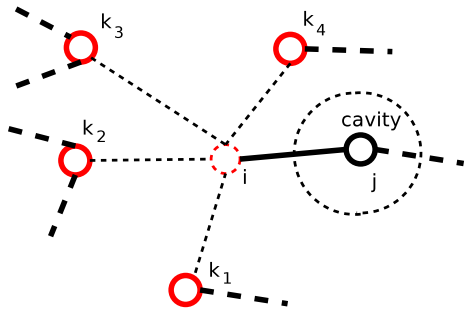


Figure 2.6: The process of iteration. We delete or add a the site i from the cavity graph $G^{(j)}$. The cavity spins are denoted in grey.

The free energy of a double site deletion is given by

$$\begin{aligned} F^{(i),(j)} &= -\beta^{-1} \sum_{\sigma_{V\setminus \partial i \cup \partial j}} \exp\left(-\beta H^{(i),(j)}\right), \\ &= F^{(i),(j)}_{\text{bulk}} + F^{(i),(j)}\left(\sigma_{\partial i \cup \partial j \setminus i,j}\right), \end{aligned} \tag{2.77}$$

with $F^{(i),(j)}$ the free energy of $P^{(i),(j)}$ and $F^{(i),(j)}_{\text{bulk}}$ the free energy of $P^{(i),(j)}_{\text{c}}$. The differences in the bulk free energies are then

$$\Delta F^{(i-j)}_{\text{iter}} \equiv F^{(j)}_{\text{bulk}} - F^{(i),(j)}_{\text{bulk}}. \tag{2.78}$$

The cavity equations

We derive the self-consistent equations (2.35) using Bayesian arguments. The main idea consists out rewriting the double-site marginal $P_{i,j}$ in two different ways.

- The first way is by deleting two sites from the graph:

$$\begin{aligned} P_{i,j}(\sigma_i,\sigma_j) \sim & \sum_{\sigma_{\partial_i\setminus j},\sigma_{\partial j\setminus i}} P^{(i),(j)}_{\partial i\setminus j,\partial j\setminus i}\left(\sigma_{\partial i\setminus j},\sigma_{\partial j\setminus i}\right) \\ & \times \exp\left[-\beta\left(\Delta H^{(i-j)}_{\text{iter}} + \Delta H^{(j-i)}_{\text{iter}} - \Delta H^{(i,j)}_{\text{link}}\right)\right]. \end{aligned} \tag{2.79}$$

We assume that the joint probability $P^{(i),(j)}_{\partial i\cup\partial j\setminus i,j}$ factorizes

$$P^{(i),(j)}_{\partial i\cup\partial j\setminus i,j}\left(\sigma_{\partial i\setminus j},\sigma_{\partial j\setminus i}\right)=\prod_{\ell\in\partial i\setminus j}P^{(j)}_{\ell}(\sigma_\ell)\prod_{\ell\in\partial j\setminus i}P^{(i)}_{\ell}(\sigma_\ell), \tag{2.80}$$

to obtain for the marginal probability distribution $P_{i,j}$:

$$\begin{aligned}P_{i,j}(\sigma_i,\sigma_j)\sim&\left[\sum_{\sigma_{\partial_i\setminus j}}\prod_{\ell\in\partial i\setminus j}P^{(i)}_{\ell}\left(\sigma_\ell\right)\exp\left(-\beta\Delta H^{(i-j)}_{\text{iter}}\right)\right]\\&\times\left[\sum_{\sigma_{\partial_j\setminus i}}\prod_{\ell\in\partial j\setminus i}P^{(j)}_{\ell}\left(\sigma_\ell\right)\exp\left(-\beta\Delta H^{(j-i)}_{\text{iter}}\right)\right]\\&\times\exp\left(-\beta\Delta H^{(i,j)}_{\text{link}}\right).\end{aligned} \tag{2.81}$$

- The second approach is by deleting a link from the graph. In this case we get

$$P_{i,j}(\sigma_i,\sigma_j)\sim P^{(i,j)}\left(\sigma_i,\sigma_j\right)\exp\left(-\beta\Delta H^{(i,j)}_{\text{link}}\right). \tag{2.82}$$

We assume now

$$P^{(i,j)}\left(\sigma_i,\sigma_j\right)=P^{(i)}_{j}\left(\sigma_j\right)P^{(j)}_{i}\left(\sigma_i\right), \tag{2.83}$$

to get

$$P_{i,j}(\sigma_i,\sigma_j)\sim P^{(i)}_{j}\left(\sigma_j\right)P^{(j)}_{i}\left(\sigma_i\right)\exp\left(-\beta\Delta H^{(i,j)}_{\text{link}}\right). \tag{2.84}$$

- Identifying the two equations (2.81) and (2.84) we get

$$P^{(j)}_{i}(\sigma)\sim\sum_{\sigma_{\partial_i\setminus j}}\prod_{\ell\in\partial i\setminus j}P^{(i)}_{\ell}\left(\sigma_\ell\right)\exp\left(-\beta\Delta H^{(i-j)}_{iter}\right). \tag{2.85}$$

For binary spins we can write

$$P^{(j)}_{i}(\sigma)=\frac{\exp\left(\beta h^{(j)}_{i}\sigma\right)}{2\cosh\left(\beta h^{(j)}_{i}\right)}, \tag{2.86}$$

to get for (2.85)

$$h_i^{(j)} = \beta^{-1} \sum_{\ell \in \partial i} \operatorname{atanh} \left(\tanh \left(\beta J_{\ell i} \right) \tanh \left(\beta h_\ell^{(i)} \right) \right). \tag{2.87}$$

The equations (2.87) are *the cavity or belief propagation equations* (BP equations). These equations can be interpreted as an algorithm that propagates fields $h_{i \to j} = h_i^{(j)}$ along the edges of the given graph G:

$$h_{i \to j}^{(t)} = \beta^{-1} \sum_{\ell \in \partial i} \operatorname{atanh} \left(\tanh \left(\beta J_{\ell i} \right) \tanh \left(\beta h_{\ell \to i}^{(t-1)} \right) \right). \tag{2.88}$$

An algorithm of this type is called a message passing algorithm (MPA). We have included in (2.88) a discrete time index t counting the sweeps we have done in a parallel updating scheme. The $2|E|$ equations (2.88) contain $2|E|$ unknowns and can be solved iteratively. Once the $h_{i \to j}$ have been found the original problem of determining the marginals of the graph P_i, is solved through the relation:

$$h_i^{(t)} = \beta^{-1} \sum_{\ell \in \partial i} \operatorname{atanh} \left(\tanh \left(\beta J_{\ell i} \right) \tanh \left(\beta h_{\ell \to i}^{(t)} \right) \right), \tag{2.89}$$

with h_i defined in an analogous way as $h_i^{(j)}$ in equation (2.86). When there are external fields θ_i present, see (2.15), one gets

$$h_{i \to j}^{(t)} \to h_{i \to j}^{(t)} + \theta_i. \tag{2.90}$$

To find the ensemble averaged versions (2.35) and (2.37) one can use again the self-averaging property combined with the independence of the different random variables in equations (2.88) and (2.89) after the introduction of a cavity. The distribution of fields $W_c(h)$ and $W(h)$, equations (2.35) and (2.37), have a clear meaning as the distribution of cavity fields and the distribution of "real" fields:

$$W_c^{(t)}(h) = \frac{\sum_{i=1}^N \sum_{j \in \partial_i} \delta \left(h - h_{j \to i}^{(t)} \right)}{cN}, \tag{2.91}$$

$$W^{(t)}(h) = \sum_{i=1}^N \frac{\delta(h - h_i^{(t)})}{N}. \tag{2.92}$$

The RS order parameters are found as $W_c(h) = \lim_{t \to \infty} W_c^{(t)}(h)$ and $W(h) = \lim_{t \to \infty} W^{(t)}(h)$. The DE equations (2.35) can be found from (2.89) using the

independence of the neighbours of a central site i when we remove this site from the graph. This independence breaks down from the moment there is a loop. Therefore, as long as the loops are much larger then the equilibration time $t_{\rm eq}$ of the MPA the equations (2.35) give the correct thermodynamic behavior. For graphs drawn from the ensemble $\mathcal{G}_{\rm deg}$ this means that $t_{\rm eq} \ll \log(N)$.

The free energy within one pure phase

Using the two assumptions (2.80) and (2.83) we find the expressions

$$F^{(i)}_{\rm cavity}(\sigma_{\partial i}) = -\sum_{k\in\partial i} h^{(i)}_k \sigma_k, \tag{2.93}$$

$$F^{(i,j)}_{\rm cavity}(\sigma_i,\sigma_j) = -h^{(i)}_j\sigma_j - h^{(j)}_i\sigma_i, \tag{2.94}$$

$$F^{(i-j)}_{\rm cavity}\left(\sigma_{\partial i\setminus j}\right) = -\sum_{k\in\partial i\setminus j} h^{(i)}_k \sigma_k, \tag{2.95}$$

for the quantities defined in (2.65), (2.71) and (2.77).

For the free energy differences (2.66), (2.72) and (2.78) we find

$$\Delta F^{(i)}_{\rm site} = -\beta^{-1}\log\left(\sum_{\sigma_{\partial i}}\prod_{k\in\partial i}\exp\left(\beta h^{(i)}_k\sigma_k + \beta J_{ik}\sigma_i\sigma_k\right)\right), \tag{2.96}$$

$$\Delta F^{(i,j)}_{\rm link} = -\beta^{-1}\log\left(\sum_{\sigma_i,\sigma_j}\exp\left(\beta h^{(i)}_j\sigma_j + \beta h^{(j)}_i\sigma_i\right)\right), \tag{2.97}$$

$$\Delta F^{(i-j)}_{\rm iter} = -\beta^{-1}\log\left(\sum_{\sigma_{\partial i\setminus j}}\prod_{k\in\partial i\setminus j}\exp\left(\beta h^{(i)}_k\sigma_k + \beta J_{ik}\sigma_i\sigma_k\right)\right). \tag{2.98}$$

Equations (2.96-2.98) are obtained with manipulations similar to the ones done in subsection 2.3.1.

The total free energy F of the distribution P is given by the expression

$$\begin{aligned} F(G) &= F\left(\{h_{i\to j}\}_{(i,j)\in E}\right) \\ &= \sum_{i=1}^{N}\Delta F^{(i)}_{\rm site}\left(\{h_{j\to i}\}\right) - \sum_{(i,j)}\Delta F^{(i,j)}_{\rm link}\left(h_{i\to j},h_{j\to i}\right). \end{aligned} \tag{2.99}$$

A disorder average of the expression (2.99) leads to the RS free energy expression $f_{\rm RS}$ (2.40).

This expression (2.99) for the free energy can be found by iteratively deleting sites from the graph. As an example, the free energy difference as a consequence of the deletion of two neighbouring sites $\Delta F_{\rm site}^{(i),(j)}$, see figure 2.7, is given by

$$\begin{aligned}\Delta F_{\rm site}^{(i),(j)} &= \sum_{\sigma_{\partial i\setminus j}} \prod_{k\in\partial i\setminus j} \exp\left(\beta h_k^{(i)}\sigma_k + \beta J_{ik}\sigma_i\sigma_k\right) \\ &\quad \times \sum_{\sigma_{\partial j\setminus i}} \prod_{k\in\partial j\setminus i} \exp\left(\beta h_k^{(j)}\sigma_k + \beta J_{jk}\sigma_j\sigma_k\right) \\ &\quad \times \exp\left(\beta J_{ij}\sigma_i\sigma_j\right) \end{aligned} \tag{2.100}$$

Using the expression (2.87) one finds the equalities:

$$\begin{aligned}\Delta F_{\rm site}^{(i),(j)} &= \Delta F_{\rm site}^{(i)} + \Delta F_{\rm iter}^{(j-i)} \\ &= \Delta F_{\rm site}^{(j)} + \Delta F_{\rm iter}^{(i-j)} \\ &= \Delta F_{\rm iter}^{(i-j)} + \Delta F_{\rm iter}^{(j-i)} + \Delta F_{\rm link}^{(i,j)}. \end{aligned} \tag{2.101}$$

From this one finds

$$\Delta F_{\rm site}^{(i),(j)} = \Delta F_{\rm site}^{(i)} + \Delta F_{\rm site}^{(j)} - \Delta F_{\rm link}^{(i,j)}. \tag{2.102}$$

One can iterate the above expressions by deleting multiple sites in order to find (2.99).

Remarks

We conclude with some remarks

- We call the two conditions (2.80) and (2.83) the Bethe-Peierls condition.
- The Bethe-Peierls condition is trivially true on a tree. The main advantage of the above derivation is that we have identified two conditions, i.e (2.80) and (2.83) which the Gibbs distribution (2.10) has to fulfil so that (2.88) and (2.99) are valid. In the literature it has indeed been observed that the equations (2.88) and (2.99) are valid for a whole set of random graphs drawn from ensembles like $\mathcal{G}_{\rm e}(N,M)$, $\mathcal{G}_{\rm deg}(N,p_{\rm d})$ and $\mathcal{G}_p$ when $N\to\infty$. The

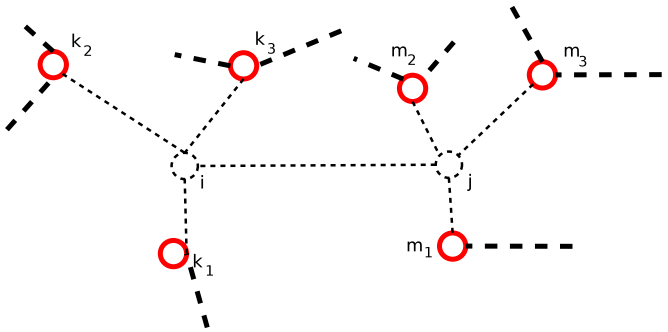


Figure 2.7: We have deleted 2 sites from the graph, i.e. the sites i and j.

underlying intuition is that these graphs have a *local tree structure*: when one removes a site from the graph the neighbours are far away and, therefore, become uncorrelated. The thermodynamic limit in this case is taken with the Bethe-Peierls condition.

- The Bethe-Peierls condition is in general not valid for Gibbs distributions of the kind (2.10). We define the set of pure states Ξ, on graphs with a local tree structure, as the minimal partition Ξ of $\mathbb{Z}_2^N$ satisfying (2.80) and (2.83), i.e.:

$$\Xi \equiv \min_n \left\{ \alpha_1 \cup \alpha_2 \cup \cdots \cup \alpha_n | \mathbb{Z}_2^N = \cup_{i=1}^n \alpha_i,\ P_{\alpha_i} \text{ satisfies Bethe} - \text{Peierls} \right\}, \tag{2.103}$$

with P_α the restriction of $P(\boldsymbol{\sigma})$ to α:

$$P_\alpha(\boldsymbol{\sigma}) \sim \begin{cases} P(\boldsymbol{\sigma}) & \text{if} \quad \boldsymbol{\sigma} \in \alpha \\ 0 & \text{if} \quad \boldsymbol{\sigma} \notin \alpha \end{cases} . \tag{2.104}$$

In a pure state α the correlations between the variables decay very fast as a function of the distance. When we remove a spin in a graph with a local tree structure the neighbouring spins will become independent within such a pure state, see figure 2.3. When the phase space can be divided in a finite number of pure states, these pure states will appear as different solutions to the cavity equations (2.88). We can reverse the reasoning and define pure states as the different solutions to the cavity equations. We are unaware of a precise connection between the solutions of the cavity equations and pure states as they are usually defined in statistical mechanics, i.e. as the extremal points of the set of solutions to the DLR equations (Dobrushin, Lanford, Ruelle) [64].

- Although the set of equations (2.88) are exact on trees the idea is to use the BP algorithm on graphs obtained from experiments such as Tanner graphs in decoding algorithms [97], networks in biology [69] or traffic networks [59]. These graphs have a finite size and can contain many loops which introduce errors in the BP algorithm. One can correct these errors using loop corrections [118].

- The set of equations (2.88) is also known in engineering and information theory as respectively BP and the sum-product algorithm [87]. In his Ph. D. thesis on low-density parity-check codes [60], Gallager used this algorithm for the first time as an efficient way to decode random sparse codes. The encoding-decoding scheme suggested by Gallager has solved the fundamental problem in coding theory, i.e. to reach the Shannon limit [93, 146].

- The cavity equations (2.87) are exact for spin models on graphs with a local tree structure. The cavity equations are valid for a broader class of spin models than the density evolution equations (2.35). An example of a graph for which equations (2.87) hold but the equations (2.35) do not hold is the Cayley tree (see figure 2.2).

2.3.2 Replica symmetry breaking in the cavity method

We first argue that local RSB considered in subsection 2.2.4 is equivalent to non-convergence of the cavity equations (2.88). Then we show how to find the 1RSB equations within the cavity method. At last we comment on global RSB.

Local symmetry breaking

At high temperatures the MPA (2.88) converges to a unique fixed point, the P solution, for $t \to \infty$. When the model contains enough frustration, the MPA does not converge anymore at low temperatures. At a certain critical temperature infinitely many solutions to the equations (2.88) appear [21]. The intuitive picture is that the initial conditions will overlap with many different solutions. Therefore, the cavity fields try to converge in different parts of the graph to different solutions.

The non-convergence of the MPA is equivalent to the non-convergence of the two-replica distribution $W_c(h^1, h^2)$, given by (2.51), to its diagonal form [6]. In [6] it is proven that the MPA converges if and only if the initial state $W_c^{(0)}(h^{(1)}, h^{(2)}) = W_c^{(\infty)}(h^{(1)})W_c^{(\infty)}(h^{(2)})$ converges to the diagonal state,
$W_c^{(\infty)}(h^{(1)}, h^{(2)}) = W_c^{\infty}(h^{(1)})\delta(h^{(1)} - h^{(2)})$, under iteration through the recursive equations (2.51) obtained through the two-replica method. The distribution

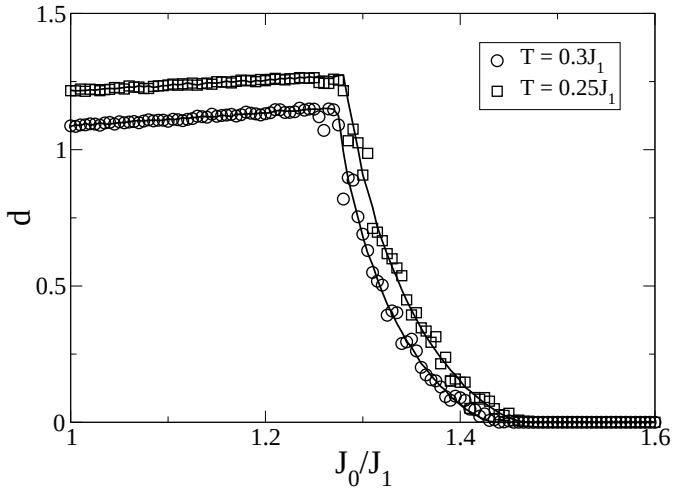


Figure 2.8: We plot the parameters d_{e}(lines) and d_{m}(markers), which quantify, respectively, the non-diagonality of the distribution in the two-replica method and the non-convergence of the cavity equations, as a function of the mean strength J_0/J_1 of the couplings of the Viana-Bray model at certain values of the temperature T. The Viana-Bray model is an Ising model on a Poissonian graph with a Gaussian distribution in the couplings. Two transition points appear, from right to left: first the system makes a transition from a ferromagnetic to a mixed phase where the messsage passing algorithm does not converge anymore. Afterwards the system becomes spin-glass like since the mean magnetization becomes zero.

$W_c^{(\infty)}(h)$ is the stable distribution of the recursive equation (2.35). In subsection 2.2.4 we have argued that this two-replica condition is equivalent to local symmetry breaking.

To study the convergence of the MPA (2.88) we define the parameter d_{m}:

$$d_{\mathrm{m}} \equiv \lim_{t\to\infty} \frac{\sum_i \sum_{j\in\partial i} \left(\tanh\left(\beta h_{i\to j}^{(t)}\right) - \tanh\left(\beta h_{i\to j}^{(t-1)}\right)\right)^2}{\sum_i \sum_{j\in\partial i}}. \tag{2.105}$$

The non-diagonality of the distribution $W_c^{(\infty)}(h^{(1)}, h^{(2)})$ is quantified through the order parameter d_{e}:

$$d_{\mathrm{e}} \equiv \int dh_1 dh_2 W_c^{(\infty)} \left(h^{(1)}, h^{(2)}\right) \left(\tanh\left(\beta h^{(1)}\right) - \tanh\left(\beta h^{(2)}\right)\right)^2. \quad (2.106)$$

In figure 2.8 we compare the two order parameters d_{m} and d_{e}. From these numerical calculations we conclude that the non-diagonality of the two-replica distribution corresponds with the non-convergence of the BP equations. The non-convergence of the MPA can also be related to the extremality or uniqueness of the Gibbs measure [167].

Multi-phase cavity method method: the first step

In the SG phase the BP algorithm can not converge to a pure state where the Bethe-Peierls assumption is correct. The reason is that the number of solutions to the set of equations (2.88) grows exponentially in the system size N [21]. It becomes hard to find a solution to the BP equations and these equations become useless. When we take a disorder average we get the RS equations (2.35) which do converge. However, the distribution $W(h)$ is a distribution of quantities $h_j^{(i)}$ that are not well-defined, see for instance in the left part of figure 2.1.

A way to deal with the non-convergence of the MPA and a number $\mathcal{O}(\exp(\sigma\mathcal{N}))$ of pure states is by defining some statistics Π on the phase space of pure states. The pure states are the solutions $\{h_{i\to j}\}_{(i,j)\in E}$ of the cavity equations, i.e.:

$$\delta_{\mathrm{cav}}\left(\{h_{i\to j}\}_{(i,j)\in E}\right) = 1, \quad (2.107)$$

with

$$\begin{aligned}\delta_{\mathrm{cav}}&\left(\{h_{i\to j}\}_{(i,j)\in E}\right) \\ &= \prod_{(i,j)\in E} \delta\left[h_{i\to j} - U\left(\{h_{k\to i}\}_{k\in\partial i\setminus j}, \{J_{ki}\}_{k\in\partial i\setminus j}\right)\right] \\ &\times \prod_{(i,j)\in E} \delta\left[h_{j\to i} - U\left(\{h_{k\to j}\}_{k\in\partial j\setminus i}, \{J_{kj}\}_{k\in\partial j\setminus i}\right)\right]. \quad (2.108)\end{aligned}$$

The free energy of the pure state $\{h_{i\to j}\}_{(i,j)\in E}$ is given by equation (2.99). After taking into consideration the form of the Gibbs measure (2.10), we define the

following statistics Π on the phase space of pure states:

$$\Pi\left(\{h_{i\to j}\}_{(i,j)\in E}\right) \sim \exp\left(-\beta m F\left(\{h_{i\to j}\}_{(i,j)\in E}\right)\right)\delta_{\rm cav}\left(\{h_{i\to j}\}_{(i,j)\in E}\right), \tag{2.109}$$

with $\beta m = \mu$. In section 2.3.1 the degrees of freedom are the spin configurations labeled by $\{\sigma_i\}_{i=1..N}$, and $\sigma_i \in \{-1,1\}$. The degrees of freedom in this section are the pure states $\left(\{h_{i\to j}\}_{(i,j)\in E}\right)$ with $h_{i\to j}\in\mathbb{R}$.

Analogously to the definition of the free energy F of the Gibbs measure (2.10), we define a generalized free energy $\Phi_{\rm 1RSB}$ ($\phi_{\rm 1RSB} = \Phi_{\rm 1RSB}/N$) and a generalized partition function $\mathcal{Z}$ of the distribution Π (2.109):

$$\begin{aligned} -\mu\phi_{\rm 1RSB}(\mu,\beta) &= \lim_{N\to\infty} N^{-1}\log\left(\mathcal{Z}\right) \\ &= \lim_{N\to\infty} N^{-1}\log\sum_{\alpha\in\Xi}\exp\left(-\mu F_\alpha(\beta)\right) \\ &= \lim_{N\to\infty} N^{-1}\log\int df\exp\left(N(\sigma(f,\beta)-\mu f)\right) \\ &= \max_f\left(\sigma(f,\beta)-\mu f\right), \end{aligned} \tag{2.110}$$

with

$$\sigma(f,\beta) = N^{-1}\log\left(\sum_\alpha \delta\left(f - f_\alpha(\beta)\right)\right). \tag{2.111}$$

Equation (2.110) should be compared with (2.7). One can find the complexity σ and the free energy f from ϕ:

$$\sigma(\mu,\beta) = \mu^2\left(\frac{\partial\phi_{\rm 1RSB}}{\partial\mu}\right)_\beta, \tag{2.112}$$

and

$$f(\mu,\beta) = \left(\frac{\partial\mu\phi_{\rm 1RSB}}{\partial\mu}\right)_\beta. \tag{2.113}$$

From (2.110) one has that the complexity decreases as a function of μ. When $\mu=\beta$ we have $\phi_{\rm 1RSB} = \overline{f}$. Due to RSB effects it is possible that $\sigma(f(\beta,\beta),\beta) < 0$.

In this case the solution of ϕ at $\mu = \beta$ is dominated by atypical pure states which disappear when $N \to \infty$. The typical value of the free energy is given by $\phi_{\rm 1RSB}(\mu^*, \beta)$, with μ^* the largest value of $\mu < 1$ such that $\sigma(\mu, \beta) \geq 0$. When $\sigma(\mu^*, \beta) = 0$ we have again $\phi(\mu^*, \beta) = f(\mu^*, \beta)$. If we assume (2.109) and (2.110) we can derive the 1RSB equations.

We follow the calculations in subsection 2.3.1 to find the marginals of the distribution Π of pure states. The variables $h_{i\to j}$ propagate along the directed edges. We use the relations (2.66), (2.72) and (2.78) when removing a site or a link. The fields $\{h_{i\to j}\}_{j\in\partial i}$ and $\{h_{j\to i}\}_{j\in\partial i}$ are kept in the probability distribution $\Pi^{(i)}$ as "external" cavity fields when removing a site i . The 1RSB cavity equations are found by writing the probability distribution Π in two different ways, analogous to subsection 2.3.1,

- First, we split the probability distribution Π into a part $\Pi^{(i),(j)}$ on the graph $G^{(i),(j)}$ (i.e., we deleted two sites) and a part containing the contributions from the i-th and the j-th site, to get

$$\begin{aligned}
&\Pi\left(\{h_{k\to l}\}_{k,l\in V}\right) \\
&\quad\sim \Pi^{(i),(j)}\left(\{h_{k\to l}\}_{k,l\in V\setminus i,j}; \{h_{i\to k}, h_{k\to i}\}_{k\in\partial i\setminus j}, \{h_{j\to k}, h_{k\to j}\}_{k\in\partial j\setminus i}\right) \\
&\quad\times \exp\left[-\mu\left(\Delta F^{(i-j)}_{\rm iter}\left(\{h_{k\to i}\}_{k\in\partial i\setminus j}\right) + \Delta F^{(j-i)}_{\rm iter}\left(\{h_{k\to j}\}_{k\in\partial j\setminus i}\right)\right)\right] \\
&\quad\times \exp\left[-\mu\Delta F^{(i,j)}_{\rm link}\left(h_{i\to j}, h_{j\to i}\right)\right] \\
&\quad\times\delta\left[h_{i\to j} - U\left(\{h_{\ell\to i}\}_{\ell\in\partial i\setminus j}, \{J_{\ell i}\}_{\ell\in\partial i\setminus j}\right)\right] \\
&\quad\times\delta\left[h_{j\to i} - U\left(\{h_{\ell\to j}\}_{\ell\in\partial j\setminus i}, \{J_{\ell j}\}_{\ell\in\partial j\setminus i}\right)\right].
\end{aligned} \tag{2.114}$$

  The distribution $\Pi^{(i),(j)}$ contains the variables $\{h_{i\to k}, h_{k\to i}\}_{k\in\partial i\setminus j}$, $\{h_{j\to k}, h_{k\to j}\}_{k\in\partial j\setminus i}$ as "cavity" variables. Next, we integrate over all variables $h_{k\to l}$, except for the two variables $h_{j\to i}$ and $h_{i\to j}$ and use the

Bethe-Peierls assumption on the marginal distribution $\Pi^{(i),(j)}_{\partial i\setminus j,\partial j\setminus i}$, to get

$$\begin{aligned}
&\Pi_{(i,j)}(h_{i\to j},h_{j\to i})\\
&= \int \prod_{k\in\partial i\setminus j} dh_{k\to j} \prod_{k\in\partial j\setminus i} \Pi^{(j)}_k\left(h_{k\to j}\right) \int \prod_{k\in\partial i\setminus j} dh_{k\to i} \prod_{k\in\partial i\setminus j} \Pi^{(i)}_k\left(h_{k\to i}\right)\\
&\quad\times \exp\left[-\mu\left(\Delta F^{(i-j)}_{\text{iter}}\left(\left\{h_{k\to i}\right\}_{k\in\partial i\setminus j}\right)+\Delta F^{(j-i)}_{\text{iter}}\left(\left\{h_{k\to j}\right\}_{k\in\partial j\setminus i}\right)\right)\right]\\
&\quad\times \exp\left[-\mu\Delta F^{(i,j)}_{\text{link}}\left(h_{i\to j},h_{j\to i}\right)\right]\\
&\quad\times\delta\left[h_{i\to j}-U\left(\left\{h_{\ell\to i}\right\}_{\ell\in\partial i\setminus j},\left\{J_{\ell i}\right\}_{\ell\in\partial i\setminus j}\right)\right]\\
&\quad\times\delta\left[h_{j\to i}-U\left(\left\{h_{\ell\to j}\right\}_{\ell\in\partial j\setminus i},\left\{J_{\ell j}\right\}_{\ell\in\partial j\setminus i}\right)\right]. \qquad (2.115)
\end{aligned}$$

- In the second approach we decompose Π as follow

$$\begin{aligned}
\Pi\left(\left\{h_{k\to l}\right\}_{k,l\in V}\right) &\sim \Pi^{(i,j)}\left(\left\{h_{k\to l}\right\}_{(k,l)\neq(i,j),(j,i)};h_{i\to j},h_{j\to i}\right)\\
&\quad\times\exp\left(-\mu\Delta F^{(i,j)}_{\text{link}}\right). \qquad (2.116)
\end{aligned}$$

We integrate over all variables except for $h_{j\to i}$ and $h_{i\to j}$ and use

$$\Pi^{(i,j)}_{(i,j)}\left(h_{i\to j},h_{j\to i}\right)=\Pi^{(i)}_j\left(h_{j\to i}\right)\Pi^{(j)}_i\left(h_{i\to j}\right), \qquad (2.117)$$

to get

$$\Pi_{(i,j)}(h_{i\to j},h_{j\to i}) \quad\sim\quad \Pi^{(i)}_j\left(h_{j\to i}\right)\Pi^{(j)}_i\left(i\to j\right)\exp\left(-\mu\Delta F^{(i,j)}_{\text{link}}\right). \qquad (2.118)$$

When we combine (2.115) and (2.118) we get

$$\begin{aligned}
\Pi^{(i)}_j\left(h_{j\to i}\right) = \int \prod_{k\in\partial j\setminus i} dh_{k\to j} \prod_{k\in\partial j\setminus i} \Pi^{(j)}_k\left(h_{k\to j}\right)\exp\left[-\mu\Delta F^{(j-i)}_{\text{iter}}\left(\left\{h_{k\to j}\right\}_{k\in\partial j\setminus i}\right)\right]\\
\times\delta\left[h_{j\to i}-U\left(\left\{h_{\ell\to j}\right\}_{\ell\in\partial j\setminus i},\left\{J_{\ell j}\right\}_{\ell\in\partial j\setminus i}\right)\right]. \qquad (2.119)
\end{aligned}$$

After taking an ensemble average over the ensemble $\mathcal{G}_{\rm deg}$ we find the 1RSB equations (2.55) with

$$W_c^{\rm 1RSB}\left(\Pi\right) = \frac{\sum_{i=1}^{N}\sum_{j\in\partial i}\delta\left(\Pi-\Pi_j^{(i)}\right)}{\sum_{i=1}^{N}\sum_{j\in\partial i}}. \tag{2.120}$$

The generalized free energy Φ is determined, analogously as done in subsection 2.3.1 with the free energy F. We get the expression

$$\begin{aligned}&\Phi\left(\left\{\Pi_{i\to j}\right\}_{(i,j)\in E}\right)\\ &\quad=\sum_{i=1}^{N}\Delta\Phi_{\rm site}^{(i)}\left(\left\{\Pi_{j\to i}\right\}_{j\in\partial i}\right)-\sum_{(i,j)}\Delta\Phi_{\rm link}^{(i,j)}\left(\Pi_{i\to j},\Pi_{j\to i}\right),\end{aligned} \tag{2.121}$$

with

$$\begin{aligned}&-\mu\Delta\Phi_{\rm site}^{(i)}\left(\left\{\Pi_{j\to i}\right\}_{j\in\partial i}\right)\\ &\quad=\quad\log\left[\int\left(\prod_{j\in\partial i}dh_j\Pi_{j\to i}^{(i)}(h_j)\right)\exp\left(-\mu\Delta F_{\rm site}^{(i)}\left(\left\{h_j\right\}_{j=1..k_i}\right)\right)\right],\end{aligned} \tag{2.122}$$

and

$$\begin{aligned}&-\mu\Delta\Phi_{\rm link}^{(i,j)}\\ &\quad=\log\left[\int dh_{j\to i}dh_{i\to j}\Pi_{i\to j}^{(j)}(h_{i\to j})\Pi_{j\to i}^{(i)}\left(h_{j\to i}\right)\exp\left(-\mu\Delta F_{\rm link}^{(i,j)}\left(\{h_{i\to j},h_{j\to i}\}\right)\right)\right].\end{aligned} \tag{2.123}$$

These equations reduce to the 1RSB equations (2.57) when taking an ensemble average. Many alternative derivations can be found in the literature [104, 149, 177].

To derive the 1RSB cavity equations (2.119) we have made the Bethe-Peierls conditions on the probability distribution Π. When the 1RSB cavity equations (2.119) admit infinitely many solutions for $N\to\infty$, these equations will not converge. In that case we have to repeat the procedure a second time and define a statistics on the space of distributions $\Pi_{i\to j}$. We get a second step in an iterative scheme that sometimes has to be repeated infinitely many times. In the literature one speaks of the first RSB step, second RSB step,... For some spin models the SG phase can be described with only one step in this scheme.

Global symmetry breaking

From the studies of the dynamics of several spin models [58, 119], one knows that the BP equations converge, but still lead to erroneous results on the typical behavior of the system. At a critical temperature $T_{\rm d}$ the relaxation time of the dynamics diverges. At this temperature $T_{\rm d}$ the 1RSB cavity equations (2.119) admit a non trivial solution at $m = 1$ with $\sigma(\mu = \beta, T_{\rm d}) > 0$. This solution corresponds with infinitely many "pure states" at some free energy $f > f_P$, with $f_{\rm P}$ the P free energy. We call this discontinuous transition a global symmetry breaking effect.

In [119] this transition was related to the properties of the Gibbs distribution P (2.1) through the divergence of the correlation length ℓ corresponding with the point-to-set correlations of the Gibbs measure P. The point-to-set correlations are defined as follows: One draws a configuration $\sigma^{(0)}$ from the Gibbs distribution P. The graph $G_\ell(i) = (V_\ell(i), E_\ell(i))$ is defined as the subgraph of G containing all vertices a minimum distance ℓ away from i. One defines the measure $P_{\ell,i}\left(\{\sigma_n\}_{n=1..N}\right)$

$$P_{\ell,i} \equiv \begin{cases} \exp\left(-\beta H\left(\{\sigma_n\}_{n=1..N}\right)\right)/Z_\ell & \text{if} \quad \sigma_n = \sigma_n^{(0)} \ \forall j \in G_\ell(i) \\ 0 & \text{if} \qquad \text{otherwise} \end{cases} . \tag{2.124}$$

The correlation length $\ell_{i,\epsilon}$ is then defined as

$$\ell_{i,\epsilon} \equiv \min\left\{\ell | \langle \sigma_{\rm i}^{(0)} \langle \sigma_{\rm i} \rangle_\ell \rangle_{\rm Gibbs} \leq \epsilon\right\}. \tag{2.125}$$

2.4 Models

It is useful to understand the thermodynamic behavior of some "prototype" models. The phase transitions of these models are similar to the ones that appear in more complicated models. Moreover, many of these models are FC and thus exactly solvable.

2.4.1 SK model

The first mean-field model we will consider is the SK model. The Hamiltonian is given by

$$H\left(\{\sigma_i\}_{i=1..N}\right) = -\sum_{0<i<j<N} J_{ij}\sigma_i\sigma_j, \tag{2.126}$$

with the J_{ij} i.i.d.r.v. drawn from $\mathcal{N}\left(\frac{J_0}{N}, \frac{J_1^2}{N}\right)$. In this case things simplify since we can express the free energy as a function of the first and second moments of $P(\boldsymbol{\sigma})$. The expression for the free energy becomes [134]:

$$-\beta f = -\frac{\beta J_0}{2n}\sum_{\alpha} m_\alpha^2 - \frac{(\beta J_1)^2}{4n}\sum_{\alpha\neq\beta} q_{\alpha\beta}^2 + \frac{1}{4}(\beta J_1)^2$$
$$+n^{-1}\log \sum_{\{S^\alpha\}_{\alpha=1..n}} \exp\left((\beta J_1)^2 \sum_{\alpha<\beta} q_{\alpha\beta} S^\alpha S^\beta + \beta J_0 \sum_\alpha m_\alpha S^\alpha\right) \quad (2.127)$$

with $q_{\alpha\beta} = \sum_{\boldsymbol{\sigma}\setminus\{\sigma^\alpha,\sigma^\beta\}} \sigma^\alpha\sigma^\beta P(\boldsymbol{\sigma})$ and $m_\alpha = \sum_{\boldsymbol{\sigma}\setminus\{\sigma^\alpha\}} \sigma^\alpha P(\boldsymbol{\sigma})$. It is possible to recover expression (2.127) from (2.31) using a high connectivity expansion [139] and a distribution for the couplings with variance $\mathcal{O}\left(c^{-1}\right)$.

In this case the RS solution f_{RS} of the free energy equals:

$$-\beta f_{\mathrm{RS}} = -\frac{\beta J_0}{2}m^2 + \frac{J_1^2\beta^2}{4}(1-q)^2$$
$$+\frac{1}{\sqrt{2\pi}}\int dz \exp\left(-\frac{1}{2}z^2\right)\ln\left(2\cosh\left(\beta J_0 m + \beta J_1\sqrt{q}z\right)\right), \quad (2.128)$$

with saddle point equations

$$m = \int dz \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)\tanh\left(\beta J_0 m + \beta J_1\sqrt{q}z\right), \quad (2.129)$$

$$q = \int dz \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)\tanh^2\left(\beta J_0 m + \beta J_1\sqrt{q}z\right). \quad (2.130)$$

When we apply the CLT to the equations (2.35) we get a Gaussian distribution with mean m and variance q for $c\to\infty$. We find that the SK model is universal in the sense that the phase diagram depends only on the variance J_1^2 and the mean J_0 of the distribution of couplings P_1. In chapter 3 we will come back on this point and study the different types of "universality classes" of FC mean field models.

Looking at the small fluctuations around the RS solution one gets the AT instability line:

$$\beta^{-2} = \int_{-\infty}^{+\infty}\frac{du}{\sqrt{2\pi}}\exp\left(-\frac{u^2}{2}\right)\mathrm{sech}^4\left(\beta\sqrt{q}u + \beta\tilde{J}_0 m\right). \quad (2.131)$$

An expression which can be obtained, both, with the Hessian [36] as well as with the two replica method.

The 1RSB solution is given by:

$$\begin{aligned} -\mu\phi_{\mathrm{1RSB}}(\mu,\beta) &= -\frac{\mu J_1^2}{4}\left[(\mu-\beta)q_1^2-\mu q_0^2+2\beta q_1-\beta\right] \\ &\quad -\frac{\mu J_0}{2}m^2+\log(2)-\int \mathcal{D}z_0 \log\int \mathcal{D}z_1 \cosh^{\mu/\beta}(\beta\Theta), \end{aligned} \tag{2.132}$$

with

$$\Theta=\beta\left(J_1\sqrt{q}_0 z_0+J_1\sqrt{q_1-q_0}z_1+J_0 m\right). \tag{2.133}$$

The saddle point equations for the 1RSB free energy read

$$m = \int \mathcal{D}z_0 \frac{\int \mathcal{D}z_1 \cosh^{\mu/\beta}(\beta\Theta)\tanh(\beta\Theta)}{\int \mathcal{D}z_1 \cosh^{\mu/\beta}(\beta\Theta)}, \tag{2.134}$$

$$q_0 = \int \mathcal{D}z_0 \left(\frac{\int \mathcal{D}z_1 \cosh^{\mu/\beta}(\beta\Theta)\tanh(\beta\Theta)}{\int \mathcal{D}z_1 \cosh^{\mu/\beta}(\beta\Theta)}\right)^2, \tag{2.135}$$

$$q_1 = \int \mathcal{D}z_0 \frac{\int \mathcal{D}z_1 \cosh^{\mu/\beta}(\beta\Theta)\tanh^2(\beta\Theta)}{\int \mathcal{D}z_1 \cosh^{\mu/\beta}(\beta\Theta)}. \tag{2.136}$$

As we have learned from the cavity method, ϕ_{1RSB} is not really the free energy but the generalized free energy. One finds the free energy of the corresponding pure states from $f=\frac{\partial\mu\Phi}{\partial\mu}$ and $\sigma=\mu^2\frac{\partial\Phi}{\partial\mu}$.

We need only two order parameters m and q to determine the phase diagram. The $(T/J_1, J_0/J_1)$ phase diagram of the SK-model is presented in figure 2.14. The RS ansatz is unstable in the SG phase and in a part of the F phase which is called the mixed (M) phase. The phase transition to the SG phase is continuous, i.e. the order parameter q behaves continuously. All phase transitions in figure 2.14 or continuous.

To find the thermodynamic value of the free energy f one has to maximize the 1RSB expression ϕ_{1RSB} to μ. The maximization of ϕ_{1RSB} to the order parameters (q_0, q_1, μ) is a consequence of a minus sign appearing in (2.127) when taking the limit $n\to 0$ after introduction of the 1RSB ansatz. The thermodynamic values of these parameters are continous as a function of the temperature.

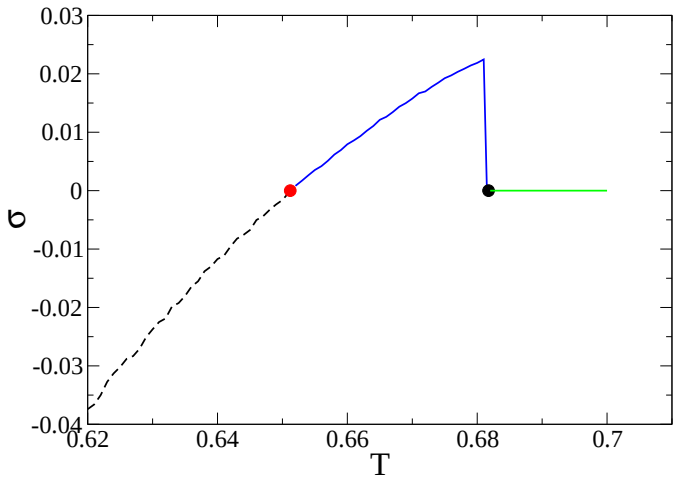


Figure 2.9: The complexity σ as a function of the temperature T of the three spin model for the Parisi RSB parameter $\mu = \beta$. The green part presents the paramagnetic solution at high temperatures. At $T_{\rm d}$ (black dot) a thermodynamic stable solution appears containing an exponential number of pure phases. This solution is presented by a blue line and becomes "unstable" at the static transition $T_{\rm c}$ (red dot). After the static transition we are left with atypical phases whose complexity is denoted by the dashed line.

The 1RSB solution given by (2.132) is unstable to 2RSB fluctuations. To find a stable solution one needs to iterate the RSB breaking procedure infinitely many times [107]. One speaks of a full RSB (FRSB) solution.

2.4.2 p-spin models

The phase diagram of p-spin models is considerably different from the phase diagram of two-spin models. The most important distinct feature in the p-spin models is the entropy crisis: the configurational entropy Σ gets zero at a certain critical temperature. The similarity with the entropy crisis, conjectured to be present in window glasses, makes the study of these models important for

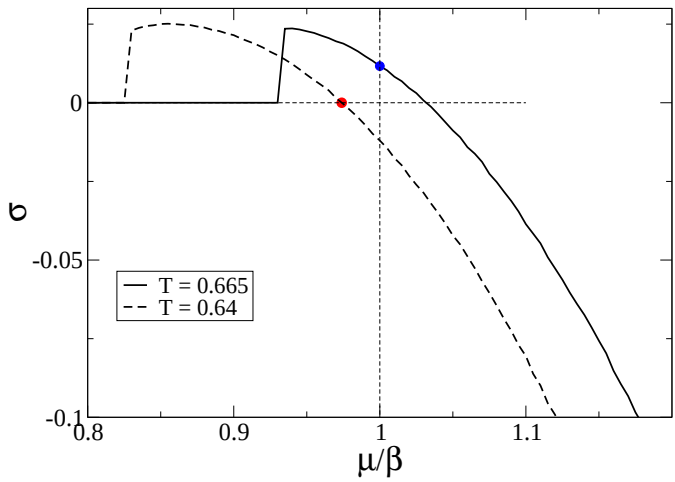


Figure 2.10: The complexity σ as a function of the RSB parameter μ of the three-spin model for a temperature $T_1 < T_c$ and $T_c < T_2 < T_d$ with $T_1/J_1 = 0.64$ and $T_2/J_1 = 0.665$. The quantities T_d and T_c denote respectively the dynamic and the static critical temperatures. The stable solutions are denoted by the filled dots. The blue dot corresponds with a SG phase containing an exponential number of pure states and the red dot with a SG phase containing a subexponential number of pure states.

structural glasses [84, 85]. We consider models with a Hamiltonian of the form

$$H\left(\{\sigma_i\}_{i=1..N}\right) = -\sum_{i_1<i_2<\cdots<i_p} J_{i_1 i_2 \cdots i_p}\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_p}. \tag{2.137}$$

The interactions $J_{i_1 i_2 \cdots i_p}$ are i.i.d.r.v. drawn from $\mathcal{N}(0, \frac{J_1^2 p!}{2N^{p-1}})$. The expression for the free energy $\phi_{1\mathrm{RSB}}$ in the 1RSB description is [37, 63, 66, 102]:

$$\phi_{1RSB} = -\frac{\beta}{4} + \frac{\mu}{4}(q_0^p - q_1^p) + \frac{\mu}{2}\left(q_1\lambda_1 - q_0\lambda_0\right) + \frac{\beta}{2}\left(\frac{q_1^p}{2} - q_1\lambda_1\right)$$

$$+\frac{\beta}{2}\lambda_1 - \frac{\ln(2)}{\beta} - \frac{1}{\mu}\int \mathcal{D}z_0 \ln \int \mathcal{D}z_1 \exp\left[-\mu f(z_0, z_1)\right],$$

with

$$f(z_0, z_1) = -\beta^{-1} \log\cosh\left(\beta\Upsilon(z_0, z_1)\right), \tag{2.138}$$

$$\Upsilon(z_0, z_1) = z_1\sqrt{\lambda_1 - \lambda_0} + z_0\sqrt{\lambda_0}, \tag{2.139}$$

$$\lambda_1 = \frac{p}{2}q_1^{p-1}, \tag{2.140}$$

$$\lambda_0 = \frac{p}{2}q_0^{p-1}. \tag{2.141}$$

The saddle point equations are

$$q_0 = \int \mathcal{D}z_0 \langle \tanh\left(\beta\Upsilon(z_0, z_1)\right)\rangle_{z_1}^2, \tag{2.142}$$

$$q_1 = \int \mathcal{D}z_0 \langle \tanh^2\left(\beta\Upsilon(z_0, z_1)\right)\rangle_{z_1}, \tag{2.143}$$

with

$$\langle f(z)\rangle_z = \frac{\int \mathcal{D}z f(z)\exp\left(-\mu f(z_0, z)\right)}{\int \mathcal{D}z \exp\left(-\mu f(z_0, z)\right)}. \tag{2.144}$$

For $\mu = 0$ one gets $\phi_{\rm RSB} = f_{\rm RS}$ and $(q_0, q_1) = (0, q)$, with q the RS order parameter. The order parameter equations contain the 1RSB solution $(q_0, \lambda_0) = (0, 0)$. We use $\sigma = \mu^2\left(\frac{\partial\phi}{\partial\mu}\right)_\beta$ and $f = \left(\frac{\partial\mu\phi}{\partial\mu}\right)_\beta$, see subsection 2.3.2, to get the complexity σ and the free energy f.

We consider three important temperatures for the p-spin models, depending on the values of the order parameters. These are the dynamical transition temperature $T_{\rm d}$, the static transition temperature $T_{\rm c}$ and the Gardner transition temperature $T_{\rm G}$, with $T_{\rm d} > T_{\rm c} > T_{\rm G}$. For a three spin model one has: $T_{\rm d}/J_1 = 0.68175(1)$, $T_{\rm c}/J_1 = 0.651(1)$ and $0.2 < T_{\rm G}/J_1 < 0.3$. The Gardner transition $T_{\rm G}$, see [63, 117] is determined by the stability of the 1RSB solution to 2RSB fluctuations. The part of the (T, μ)-space where the 1RSB ansatz is no longer stable is given by [117]:

$$\frac{2}{p(p-1)\beta^2 q_1^{p-2}} < \frac{\int \mathcal{D}z \cosh^{\mu/\beta-4}\beta\lambda_1 z}{\int \mathcal{D}z \cosh^{\mu/\beta}\left(\beta\lambda_1 z\right)}. \tag{2.145}$$

The order parameters have the following behavior:

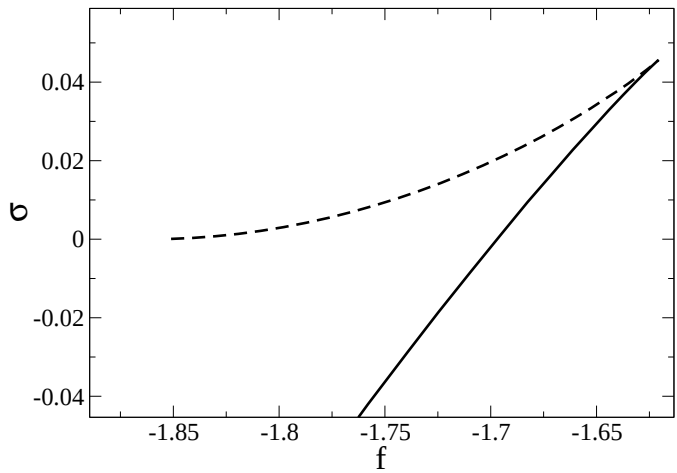


Figure 2.11: The complexity σ of the three spin model as a function of the free energy f of the "pure" phases for $T = 0.3$. The convex part (dashed line) consists of a non-physical solution and corresponds with a complexity which increases as a function of μ (it corresponds with the dashed part in the figure 2.12). Only the concave part (solid line) of the complexity is physical.

- $T > T_{\rm d}$: The P solution $(q_0, q_1) = 0$ is the only solution of the order parameter equations for all values of μ. The P solution is colored gree in figure 2.9.

- $T_{\rm d} > T > T_{\rm c}$: The order parameter equations have two stable solutions (to 2RSB fluctuations). A P solution and a non-trivial solution appearing at $\mu = \beta$. This solution has a positive complexity $\sigma(\mu = \beta, \beta) > 0$ and a generalized free energy $\phi_{\rm 1RSB} = f_{\rm P}$. The phase consists of an entropic amount $\mathcal{N} \sim \exp(N\sigma(f, \beta))$ of phases with a free energy f. This solution is denoted as a blue line/dot in figures 2.9 and 2.10.

- $T_{\rm c} > T > T_{\rm G}$: The complexity is negative for $\mu = \beta$. The thermodynamic value is given by a point $\mu_c < \beta$ where $\sigma(\mu_c) = 0$ and $\phi_{\rm 1RSB} > f_{\rm P}$. This phase consists of a subexponential amount of phases of free energy f. The vanishing of the complexity $\sigma(f, \beta)$, also sometimes called the configurational entropy, corresponds with a condensation transition in the phase space of pure states. This solution is shown as a red dot in figures 2.9 and 2.10. The

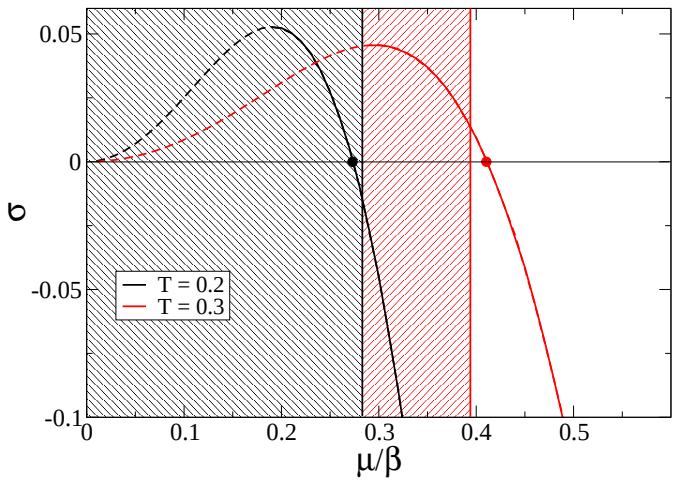


Figure 2.12: The complexity σ for the three-spin model as a function of the Parisi RSB parameter μ/β. The thermodynamic point is denoted with a filled dot while the region that is unstable to 2RSB fluctuations is shaded. The Gardner transition $T_{\rm G}$ is located between $T/J_1 = 0.2$ and $T/J_1 = 0.3$. The dashed part of the complexity is unphysical: from the relation (2.110) we have that the complexity must decrease as a function of μ. It corresponds through a Legendre-Fenchel transform with the dashed part in figure 2.11.

1RSB ansatz is stable to 2RSB fluctuations in this region. In figure 2.12 the unstable part of the complexity curve is shaded. For $T = (0.3J_1) > T_G$, the complexity is still positive in a thermodynamic stable region (red dot).

- $T < T_{\rm G}$: The 1RSB ansatz is unstable to 2RSB fluctuations. In figure 2.12 one notices that for $T = 0.2$ the 1RSB ansatz is unstable. The exact solution in this region is a FRSB solution.

The thermodynamic value of the order parameter q makes a discontinuous jump while crossing the thermodynamic transition. We have a discontinuous transition. Using a Legendre transform one finds the complexity as a function of the free energy f: in figure 2.11 the complexity contains two branches: a non-physical convex branch and a physical concave branch. The non-physical branch comes from the part of the complexity that increases as a function of μ in figure (2.12).

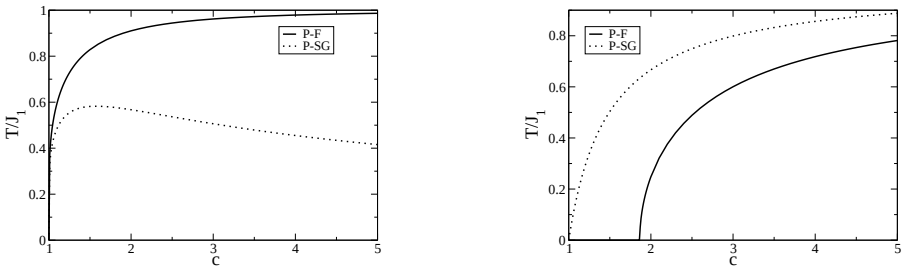


Figure 2.13: The critical temperatures T/J_1 as a function of the mean connectivity c for the bimodal Viana Bray model (left) and the Gaussian Viana Bray model (right). In the Gaussian model $J_0 = J_1 = 1$ while in the bimodal model $\rho = 1$. We remark that the P-SG transition is ρ independent for the bimodal model.

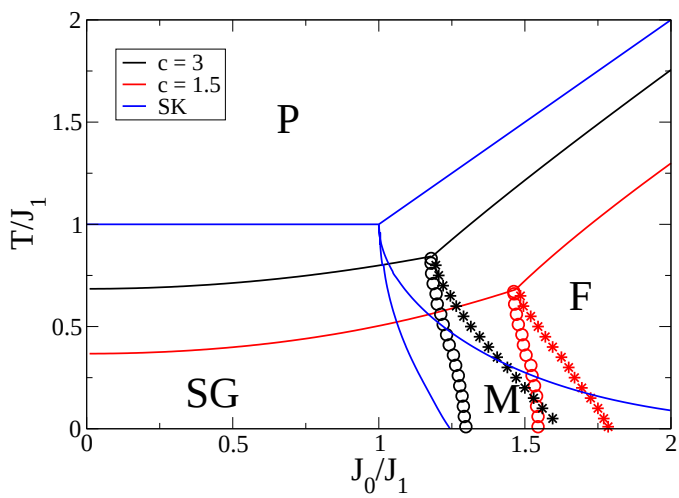


Figure 2.14: The $(T/J_1, J_0/J_1)$-phase diagram of the Viana-Bray model for Gaussian couplings with given values of the mean connectivity c. Dots denote the M to SG transition and stars denote the M to F transition. The blue lines denote the phase diagram of the SK model.

2.4.3 Viana-Bray model

In the literature one usually refers to the Viana-Bray model as a FiC model with $p_{\rm d}(k) = p_{\rm p}(k;c)$. For FiC models the distribution of effective fields is not Gaussian. At low temperatures the distribution becomes a sum of delta functions with a continuous background, see figure 2.1. The presence of the continuous background for $T \to 0$ indicates the instability to RSB [105]. For the distribution of couplings $P_{\rm I}$ we consider bimodal or Gaussian distributions:

- the bimodal model:

$$P_B(J) \equiv \left(\frac{1+\rho}{2}\right)\delta\left(J - J_1\right) + \left(\frac{1-\rho}{2}\right)\delta\left(J + J_1\right), \tag{2.146}$$

- the Gaussian model:

$$P_G(J) \equiv \mathcal{N}\left(\frac{J_0}{c}, cJ_1^2\right). \tag{2.147}$$

The P to F and P to SG phase transitions are denoted as a function of the mean connectivity in figures 2.13. All phase transitions dissapear at $c = 1$, which is the percolation threshold for a Poissonian random graph. When the largest component in the graph is not extensive phases do not appear. In the Gaussian model we see a continuous increase of the critical temperatures to the fully-connected case. In figure 2.14 we present the phase diagrams for different values of c for the Viana-Bray model with Gaussian couplings. We notice a decrease of the SG phase and the M phase due to the finite connectivity. One can understand this result from the decrease of frustration as a result of less frustrated loops when the connectivity decreases.

2.4.4 Random-energy model

The random energy model (REM) was introduced by Derrida in [38]. The p-spin model with random couplings converges to the random energy limit when $p \to \infty$. The random energy model can be defined through:

$$P(\{\sigma_i\}_{i=1..N}) \equiv \exp\left[E\left(\{\sigma_i\}_{i=1..N}\right)\right], \tag{2.148}$$

with $E\left(\{\sigma_i\}_{i=1..N}\right)$ i.i.d.r.v. drawn from some distribution $P(E) = \mathcal{N}\left(0, NJ^2\right)$. The configurations $\{\sigma_i\}_{i=1..N}$ are also called the states or the levels. The simplicity of the REM makes the model a first attempt in understanding the behavior of complex systems. The REM gives insights in the nature of the

freezing/condensation transition in glassy systems. The freezing transition of a random heteropolymer into a frozen globule is similar to the P-SG transition found in the REM [10, 158]. In chapter 3 we will use the REM to analyze "random coding".

We define the entropy $s(e)$, through the average number of states $\overline{\mathcal{N}}(E) = \exp[Ns(E)]$ which have an energy in the interval $[E, E + dE[$. The entropy s equals, for $N \to \infty$,

$$s(\epsilon) = \log(2) - \frac{\epsilon^2}{2J^2}. \tag{2.149}$$

As a consequence, we have three regions for $N \to \infty$, depending on the value ϵ with respect to $\epsilon_c = \sqrt{2\log(2)}J$:

- $\epsilon > \epsilon_c$: the average number of states increases exponentially in the system size
- $\epsilon = \epsilon_c$: the number of states scales subexponentially in the system size
- $\epsilon < \epsilon_c$: for typical configurations there are almost no energy levels with these energies, for $N \to \infty$

In the canonical ensemble, i.e., if we use the Legendre-Fenchel transformation $f(T) = \max_e(e - s(e)T)$ we get $s(T) = \ln(2) - \frac{(J)^2}{4T^2}$. At low temperatures, $T < T_c = J/(2\sqrt{\ln(2)})$, the average of the free energy is concentrated on atypical states. Since we are interested in the typical behavior we set

$$s = \begin{cases} \ln(2) - \frac{(J)^2}{4T^2} & \text{when } T > T_c \\ 0 & T \leq T_c \end{cases} \tag{2.150}$$

The interpretation is clear: below the critical temperature the system freezes or condensates in a subexponential number of states [102]. The order parameter of this condensation transition is given by the participation ratio $\overline{Y}$ defined through

$$\overline{Y} \equiv \overline{\sum_{\{\sigma_i\}_{i=1..N}} \left(\frac{\exp\left[-\beta E\left(\{\sigma_i\}_{i=1..N}\right)\right]}{\sum_{\{\sigma_i\}_{i=1..N}} \exp\left[-\beta E\left(\{\sigma_i\}_{i=1..N}\right)\right]} \right)^2}. \tag{2.151}$$

When $\overline{Y} = 0$ the thermodynamic weight is spread over an exponential number of states while for $\overline{Y} > 0$ the thermodynamic weight is concentrated on a

subexponential number of states. For the REM it is equal to

$$\overline{Y} = \begin{cases} 0 & \text{when } T > T_c \\ 1 - \frac{T}{T_c} & \text{when } T < T_c \end{cases} \tag{2.152}$$

It is instructive to consider how the condensation transition appears within the replica theory. Within the RS ansatz $q = 0$ and $-\beta f_{\mathrm{RS}}(\beta) = \frac{\beta^2}{4} + \ln(2)$. There is only a P phase and no freezing transition. However, within the 1RSB ansatz there, is below the critical temperature T_c, the SG solution with $q_0 = 0$, $q_1 = 1$ and $m = T/T_c$. The free energy of this phase is temperature independent $-\beta f_{1\mathrm{RSB}} = \frac{\beta\beta_c}{4} + \frac{\beta}{\beta_c}\ln(2)$ and the entropy is zero.

2.5 Summary

The replica method and the cavity method are two well-developed methods to study the statistical mechanics of FiC systems. The replica method within the replica-symmetric approximation leads to a description of the free energy as a function of a distribution $W_c(h)$. This distribution can be identified with the distribution of cavity fields which are found as solutions to the belief-propagation equations. The belief-propagation equations are equations in the marginals of the Gibbs distributions on a given graph instance. The replica symmetric description is valid for Gibbs distributions on trees that fulfill the Bethe-Peierls conditions. In fact, the Bethe-Peierls conditions must be invalid in a very "extreme way" in order to have that the RS description is invalid. This happens in the SG-phase where the phase space must be divided in infinitely many pure states.

The correct values of thermodynamic quantities in the SG-phase can be found through an iterative scheme. In the replica theory one makes a specific ansatz that breaks the replica symmetry. In the cavity method one defines a statistics over the phase space of pure states. One considers three types of spin glass phases in disordered systems depending on the nature of the order parameter:

- A 1RSB spin glass phase: This phase transition appears in p-spin models. The SG-phase is correctly described by one step in the iterative scheme. One finds a thermodynamic stable solution with an exponential number of pure states.
- A FRSB spin glass phase: In this spin glass phase one has to break the replica symmetry an infinite number of times. The SG phase in the SK-model and in fully-connected p-spin models below the Gardner transition are examples. This phase consists of a subexponential number of pure states.

- A condensation transition: in this case one can describe the phase transition with one step in the iterative scheme. The entropy gets negative at the transition and the partition function is dominated by a subexponential number of spin configurations.

Chapter 3

Lévy spin glasses

The phase diagram of a spin model on a random graph with i.i.d.r. couplings with finite variance converges to the phase diagram of the Sherrington-Kirkpatrick model when the mean connectivity becomes infinitely large. This is a central limit theorem for spin glasses. The SK-model is universal in the sense that the phase diagram does not depend on the precise details of the coupling distribution. However, for distributions with a divergent variance in the couplings, spin models on graphs do not converge to the SK model. Stable distributions with a divergent variance are called Lévy distributions. Such distributions appear in models for spin glasses [86] but also in models for the correlations of stocks on the stock market [22].

We study a generalization of the standard mean-field model for spin glasses, i.e. the SK model, to a fully-connected mean-field model where the couplings are drawn from Lévy distributions. Such a model has been introduced in [28] and the phase diagram has been derived using the central limit theorem. The authors of this paper have determined the phase diagram assuming a Gaussian distribution for the distribution of cavity fields. Using this assumption they have found a spin-glass phase consisting of one pure phase. The spin glass phase of the SK-model contains infinitely many pure states. It has been argued that the reason for this curcial difference lies in the presence of strong bonds in the Lévy spin glass.

In [79] it has been shown that the Gaussian assumption on the distribution of cavity fields is wrong. We show how to determine without the Gaussian assumption the macroscopic quantities of disordered systems with the disorder drawn from a distribution with an infinitely large variance. The main ingredient of our solution consists of decomposing the system in a finitely connected system of strong bonds and an extensively connected system of weak bonds. We derive the phase diagram of Lévy spin glasses through this decomposition using three different procedures: the replica method, the cavity method and simulations. The resultant phase diagram

is universal: the phase diagram of spin models on graphs with a divergent variance in the couplings converges to this one when the mean connectivity diverges. We find a spin-glass phase similar to the spin-glass phase of the SK-model, which is different than the conjectures made in [28]. Similar ideas have been published in [78]. Recently we have applied a similar procedure to the study of Lévy random matrices [30]. We have decomposed random matrices in a sparse random matrix containing the strong matrix elements and a dense matrix containing the weak matrix elements. Using this procedure we have determined the spectrum and the localization properties of eigenvectors of Lévy random matrices [101].

3.1 Motivation

3.1.1 The SK model and the CLT

The standard mean-field model for spin glasses is the SK model [83, 161]. Besides the original interests in the study of glassy systems this model has been important in the modelling of problems ranging from biology, computer science, artificial neural networks to economics (see chapter 1). Essential in the study of the SK model at high temperatures is the "standard" CLT, see section 2.4.1. Moreover, because of the CLT this model contains universality. All coupling distributions that satisfy the CLT lead in the fully-connected (FC) case to the SK solution with the appropriate mean J_0 and variance J_1^2 in the coupling distribution of (2.126).

3.1.2 Lévy distributions in spin glasses

An example of a system where Lévy distributions appear are spin glasses. Spin glasses are magnetic moments doped in a noble metal. These magnetic moments interact indirectly through the polarization of the conduction electrons of the metal. This interaction is described by the RKKY-theory [151] with coupling strength:

$$J(r) \sim \frac{r\cos(r) - \sin(r)}{r^4}, \tag{3.1}$$

with r the distance between two magnetic moments.

In [86] the effective field distribution for RKKY-spin glasses has been studied. This distribution turned out to be a Cauchy distribution for small concentrations of the diluted moments. This is in contrast with a Gaussian distribution predicted by the SK model. A more accurate numerical investigation in [12] shows that there is a transition as a function of the concentration of diluted moments: for low concentration values the distribution is Cauchy while for higher values it becomes

Gaussian. Inspired by [86] a new class of mean field spin glasses was introduced by Cizeau and Bouchaud [28]. This model is a FC model where the interactions are drawn from Lévy distributions. In [28] the model has been studied using the cavity method within a Gaussian assumption for the distribution of fields. According to this analysis the spin glass phase is RS at the phase transition. The phase space of RS spin glasses contains one pure phase, see figure 3.1. The authors of [28] conclude that the SG phase in Lévy spin glasses is strikingly different from the SG phase of the SK model which contains infinitely many pure states, see figure 3.1.

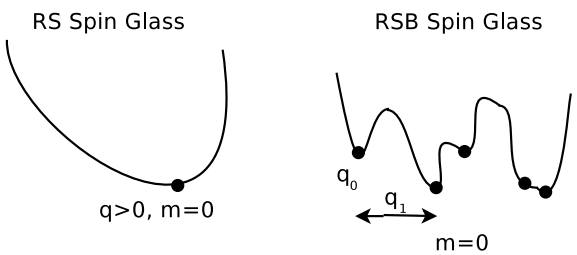


Figure 3.1: Left: The free energy landscape of a replica symmetric spin glass. Right: The free energy landscape for a spin glass where the replica symmetry is broken.

The aim of our work is to test the different conjectures about the nature of the spin glass phase made in [28] and to derive the complete phase diagrams of the model without the Gaussian assumption. Recently it has been shown that this assumption is wrong [79]. We would like to develop methods to determine the macroscopic behaviour of disordered systems where the interactions have a broad range. These problems appear difficult since one can not use the CLT. We use the methods of section 2, i.e. the replica method and the cavity method.

3.2 Definitions and notations

3.2.1 The Hamiltonian

The model we consider is defined through:

$$H(\{\sigma_i\}_{i=1..N}) = -\frac{1}{2N^{1/\alpha}} \sum_{i \neq j} J_{ij}\sigma_i\sigma_j, \tag{3.2}$$

with the couplings J_{ij} i.i.d.r.v. drawn from a stable or equivalently Lévy distribution. The scaling with $N^{1/\alpha}$ is such that $H \sim \mathcal{O}(N)$. The sum of N i.i.d.r.v. random variables drawn from a Lévy distribution scales as follow:

$$\sum_{n=1}^{N} X_n = N^{1/\alpha} X + d, \tag{3.3}$$

with X distributed according to the same N-independent distribution as the X_n variables. The fields

$$h_i^{\text{exch}} \equiv \sum_j \frac{J_{ij}}{N^{1/\alpha}} \sigma_j \ , \tag{3.4}$$

are thus random variables of order $\mathcal{O}(1)$. Important in this model is the appearance of a finite number of strong bonds of order $\mathcal{O}(1)$ and an infinite number of weak bonds of order $\mathcal{O}(N^{-1/\alpha})$. Indeed, the large tail behaviour of a Lévy distribution is given by

$$P(J) \sim \frac{1}{|J|^{\alpha+1}}, \tag{3.5}$$

for $\alpha < 2$. Therefore, if we define a strong bond through $|J| > N^{1/\alpha}$, the probability to have a strong bond is proportional to:

$$\int_{N^{1/\alpha}}^{+\infty} \frac{1}{J^{\alpha+1}} \sim \frac{1}{N\alpha}. \tag{3.6}$$

According to the above definition of a strong bond, the field h_i^{exch} contains a contribution from a finite number of strong bonds J_{ij}.

3.2.2 Lévy distributions

We define the stable distributions $P_\alpha^{J_1,\gamma,J_0}(J)$ through their characteristic function $L_\alpha^{J_1,\gamma,J_0}(q)$:

$$P_\alpha^{J_1,\gamma,J_0}(J) \equiv \int \frac{dq}{2\pi} \exp\left(-iqJ\right) L_\alpha^{J_1,\gamma,J_0}(q). \tag{3.7}$$

The characterstic function is of the form

$$L_\alpha^{J_1,\gamma,J_0}(q) = \exp\left[i\frac{qJ_0}{N^{1-1/\alpha}} - \left|\frac{J_1 q}{\sqrt{2}}\right|^\alpha \left(1 - i\gamma\Phi \text{sign}(q)\right)\right]. \tag{3.8}$$

The distribution $P_\alpha^{J_1,\gamma,J_0}(J)$ is characterized by four parameters: the exponent $\alpha \in (0,1) \cup (1,2]$, the skewness $\gamma \in [-1,1]$, the scale parameter $J_1 > 0$ and the shift $J_0 \in \mathbb{R}$. The quantity Φ is given by $\Phi = \tan\left(\frac{\alpha\pi}{2}\right)$. Lévy distributions contain two different parameters that control the bias in the couplings: J_0 and γ. We refer the reader to appendix A for a discussion of the role of α and γ. For $\alpha = 1$ and $\gamma \neq 0$ the quantity Φ has a different expression and we do not consider this case in the sequel.

The SK model is obtained for $\alpha = 2$ independent of γ: for $\alpha = 2$, $P_\alpha^{J_1,\gamma,J_0}(J)$ is a Gaussian distribution with mean J_0/N and variance J_1^2/N [161]. For $\alpha < 2$ and $-1 < \gamma < 1$, the asymptotic behavior $\rho(J)$ of $P_\alpha^{J_1,\gamma,J_0}(J)$ for $|J| \to \infty$ can be derived from the explicit form of $L_\alpha^{J_1,\gamma,J_0}(q)$:

$$\rho(J) \equiv \lim_{|J|\to\infty} P_\alpha^{J_1,\gamma,J_0}(J) = (1 + \gamma\,\mathrm{sign}J)\,\frac{C_\alpha}{|J|^{\alpha+1}}, \tag{3.9}$$

where

$$C_\alpha = \left(\frac{J_1}{\sqrt{2}}\right)^\alpha \frac{1}{\pi}\sin\left(\frac{\alpha\pi}{2}\right)\Gamma(\alpha+1). \tag{3.10}$$

Accordingly, the integrals for the second and higher moments of the distribution diverge for $\alpha < 2$ due to the power-law decay in equation (3.9).

We can define stable distributions through (3.8) without losing any generality [135, 140]. We remark that there are many equivalent definitions possible for the characterstic function L [135]. Loosely speaking, a random variable x is stable if the sum of $n > 1$ independent and identical copies of x is characterized by the same distribution as the original variable, exhibiting only a different scale and shift.

3.3 The cavity method within the Gaussian assumption

In [28] Cizeau and Bouchaud have studied the Lévy spin glass using a Gaussian assumption for the distribution of fields. The cavity method, see section 2.3, leads, within the one pure phase assumption, directly to the following equations (see equation (2.87)):

$$h_j = \beta^{-1}\sum_{i=1}^{N} \mathrm{atanh}\left(\tanh(\beta h_i)\tanh\left(\beta\frac{J_{ij}}{N^{1/\alpha}}\right)\right). \tag{3.11}$$

Equation (3.11) contains a sum of N i.i.d.r.v. with a finite variance. Therefore, as argued in [28], one can apply the CLT. As long as $\gamma = 0$ and $J_0 = 0$ one can solve

(3.11) for $N \to \infty$ using a Gaussian assumption for the distribution of the cavity fields W:

$$W(h) = \frac{1}{\sqrt{2\pi Q}} \exp\left(-\frac{h^2}{2Q}\right). \tag{3.12}$$

One gets for the variance Q the selfconsistent equation

$$Q = \frac{C_\alpha}{\beta^2} \int dh W(h) \int dJ \frac{1}{|J|^{\alpha+1}} \text{atanh}^2 \left(\tanh(\beta h) \tanh\left(\beta J\right)\right). \tag{3.13}$$

Performing a stability analysis one finds, for the critical P-SG value β_{SG}:

$$1 = C_\alpha \int \frac{dJ}{|J|^{\alpha+1}} \tanh^2\left(\beta_{\text{SG}} J\right). \tag{3.14}$$

3.4 The replica theory

The cavity method in the previous subsection assumes a Gaussian distribution for the cavity fields. We follow a replica calculation which does not make such an assumption [79].

3.4.1 The distribution of effective fields

To calculate the disorder average over the quenched couplings J_{ij} of the free energy corresponding with the Hamiltonian (3.2) we follow the replica method, see subsection 2.2. However, for Lévy spin glasses there is a complication since the integer moments $\overline{Z^n}$ of the partition function diverge for real β due to the power-law behavior of $P_\alpha^{J_1,\gamma,J_0}(J)$ for $|J| \to \infty$. As noted in reference [79], the introduction of an imaginary temperature $\beta = -i\tilde{\beta}$, with a real parameter $\tilde{\beta} > 0$, allows a straightforward calculation of the average $\overline{Z^n}$ by means of the definition of the characteristic function, equation (3.8). Because the distribution is not Gaussian, the calculations in [79] are similar to the ones in subsection 2.2. Following these calculations we find the expression for the averaged free energy $\overline{f}$ in the limit $N \to \infty$:

$$\overline{f} = \overline{f}_1 + \overline{f}_2, \tag{3.15}$$

with

$$i\tilde{\beta}\overline{f}_1 = -\lim_{n\to 0}\frac{1}{2n}\sum_{\boldsymbol{\sigma}\boldsymbol{\tau}}\Bigg(-\Big(\frac{J_1\tilde{\beta}}{\sqrt{2}}\Big)^{\alpha}|\boldsymbol{\sigma}.\boldsymbol{\tau}|^{\alpha}\left(1+i\gamma\mathrm{sign}(\boldsymbol{\sigma}\cdot\boldsymbol{\tau})\Phi\right) \\ -i\tilde{\beta}J_0(\boldsymbol{\sigma}.\boldsymbol{\tau})\Bigg)P(\boldsymbol{\sigma})P(\boldsymbol{\tau})\ , \tag{3.16}$$

$$i\tilde{\beta}\overline{f}_2 = \lim_{n\to 0}\frac{1}{n}\log\Bigg\{\sum_{\boldsymbol{\sigma}}\exp\Bigg[-\sum_{\boldsymbol{\tau}}\Bigg(-\Big(\frac{J_1\tilde{\beta}}{\sqrt{2}}\Big)^{\alpha}|\boldsymbol{\sigma}.\boldsymbol{\tau}|^{\alpha}\left(1+i\gamma\mathrm{sign}(\boldsymbol{\sigma}\cdot\boldsymbol{\tau})\Phi\right) \\ -i\tilde{\beta}J_0(\boldsymbol{\sigma}.\boldsymbol{\tau})\Bigg)P(\boldsymbol{\tau})\Bigg]\Bigg\}. \tag{3.17}$$

The order parameter $P(\boldsymbol{\sigma})$, with $\boldsymbol{\sigma} = \left(\sigma^1,\sigma^2,\ldots,\sigma^n\right)$, fulfills the self-consistent equation

$$P(\boldsymbol{\sigma}) \quad = \frac{\exp\Big(\sum_{\boldsymbol{\tau}}P(\boldsymbol{\tau})\Big[-i\tilde{\beta}J_0\boldsymbol{\sigma}\cdot\boldsymbol{\tau}-\Big|\frac{J_1\tilde{\beta}\boldsymbol{\sigma}.\boldsymbol{\tau}}{\sqrt{2}}\Big|^{\alpha}(1+i\gamma\mathrm{sign}(\boldsymbol{\sigma}\cdot\boldsymbol{\tau})\Phi)\Big]\Big)}{\sum_{\boldsymbol{\sigma}}\exp\Big(\sum_{\boldsymbol{\tau}}P(\boldsymbol{\tau})\Big[-i\tilde{\beta}J_0\boldsymbol{\sigma}\cdot\boldsymbol{\tau}-\Big|\frac{J_1\tilde{\beta}\boldsymbol{\sigma}.\boldsymbol{\tau}}{\sqrt{2}}\Big|^{\alpha}(1+i\gamma\mathrm{sign}(\boldsymbol{\sigma}\cdot\boldsymbol{\tau})\Phi)\Big]\Big)}. \tag{3.18}$$

Analogous to the ansatz made in equation (2.33) for the order parameter of FiC, we make the following RS ansatz,

$$P(\boldsymbol{\sigma}) = \int dhW(h)\prod_{a=1}^{n}\frac{\exp\left(-i\tilde{\beta}h\sigma^a\right)}{2\cosh\left(-i\tilde{\beta}h\sigma^a\right)}, \tag{3.19}$$

which defines the field distribution $W(h)$. We make an analytic continuation of the temperature to real values by taking $\tilde{\beta} = i\beta$ at the end of the calculation. Substitution of (3.19) into (3.18) gives:

$$W(h) = \int\frac{ds}{2\pi}\exp\left[ish\right]\exp\Bigg[-\int dhW(h)\int\frac{d\hat{J}dJ}{2\pi} \\ \times\left(\Big(\frac{J_1}{\sqrt{2}}\Big)^{\alpha}|\hat{J}|^{\alpha}\Big(1+i\gamma\Phi\mathrm{sign}(\hat{J})\Big)+iJ_0\hat{J}\right)\exp\Big[i\hat{J}J\Big]\ f(J,h,s)\Bigg]. \tag{3.20}$$

The function $f(J,h,s)$ is defined as

$$f(J,h,s)\equiv\exp\left(-\frac{is}{\beta}\mathrm{atanh}\Big[\tanh\left(\beta J\right)\tanh\left(\beta h\right)\Big]\right). \tag{3.21}$$

We integrate in equation (3.20) over the $\hat{J}$ variable, using the expressions in appendix C.1, to obtain the following simplified expression for $W(h)$

$0 < \alpha < 1$:

$$W(h) = \int \frac{ds}{2\pi} \exp\left[ish - isJ_0 m + \int dh W(h) \int_{-\infty}^{\infty} dJ \rho(J) \Big(f(J,h,s) - 1 \Big)\right], \tag{3.22}$$

$1 < \alpha < 2$:

$$W(h) = \int \frac{ds}{2\pi} \exp\Big(ish - isJ_0 m\Big) \times \exp\left[\int dh W(h) \int_{-\infty}^{\infty} dJ \rho(J) \Big(f(J,h,s) - f'(0,h,s)J - 1\Big)\right], \tag{3.23}$$

where $f'(0,h,s) = \frac{\partial f(J,h,s)}{\partial J}\Big|_{J=0}$. The distribution of couplings $\rho(J)$ is defined by equation (3.9). The RS magnetization m and the RS spin-glass order-parameter q are determined through the averages

$$m = \int dh W(h) \tanh(\beta h), \qquad q = \int dh W(h) \tanh^2(\beta h). \tag{3.24}$$

Only the large tail behavior of the distribution $P_\alpha^{J_1,\gamma J_0}$ appears in the equations (3.22) and (3.23). This could mean that the system exhibits a certain degree of universality, i.e. the thermodynamic behavior only depends on the large tail behavior of the coupling distribution $P(J)$.

The distribution $\rho(J)$ is symmetric when $\gamma = 0$, with equations (3.22) and (3.23) reducing to a single equation, obtained previously in [79]. The equations (3.22) and (3.23) do not have a Gaussian distribution for $\alpha < 2$.

3.4.2 The normalization of the coupling distribution through a cutoff

The main difficulty in equations (3.22) and (3.23) concerns the normalization of $\rho(J)$ since the integral $\int dJ \rho(J)$ diverges for $\alpha < 2$. Therefore, it is not possible to normalize the distribution. This invalidates the numerical calculation of $W(h)$

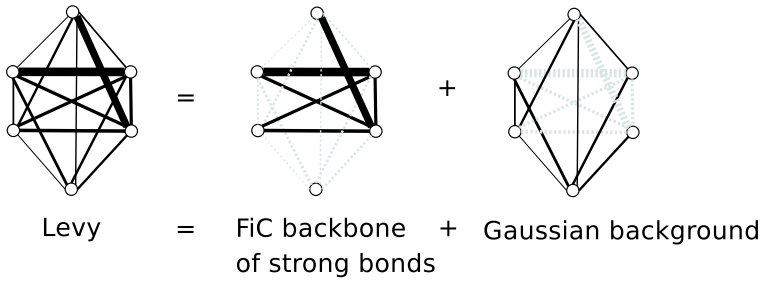


Figure 3.2: The Lévy spin glass is studied through an explicit decomposition of the bonds in a set of strong bonds and a set of weak bonds through the choice of a certain cutoff value. The width of the bonds is proportional to the couplings strength drawn from a Lévy distribution with $\gamma = 0$, $\alpha = 1$. For large system sizes the strong bonds form a FiC backbone where the couplings are drawn according to the large tail behavior of the Lévy distribution. The weak bonds are approximatively drawn from a Gaussian distribution with a mean and a variance which depend on the Lévy distribution.

through the population dynamics algorithm [104] because it is not possible to sample random numbers from a non-normalizable distribution.

In this subsection we propose a simple procedure that allows to normalize $\rho(J)$ and to derive a self-consistent equation for $W(h)$ which is similar to the order-parameter equation of FiC on random graphs equation (2.35). The numerical solution of this equation can be obtained through the population dynamics algorithm.

The method consists of making an explicit separation between strong and weak bonds, see figure 3.2. This is realized through the insertion of a temperature dependent cutoff $T\epsilon > 0$ in the integrals over J occurring in equations (3.22) and (3.23), splitting each of them into an integral around zero (from $-T\epsilon$ to $T\epsilon$) plus an integral over the couplings that satisfy $|J| > T\epsilon$. Assuming $T\epsilon \ll 1$, the integrals around zero can be performed analytically by expanding $f(J, h, s)$ around $J = 0$

up to order $\mathcal{O}(J^2)$, resulting in the following equations

$$0<\alpha<1:$$

$$\int_{-\infty}^{\infty} dJ\rho(J)\Big[f(J,h,s)-1\Big] = -2is\gamma C_\alpha \tanh(\beta h)\frac{(T\epsilon)^{1-\alpha}}{1-\alpha} - s^2 C_\alpha \tanh^2(\beta h)\frac{(T\epsilon)^{2-\alpha}}{2-\alpha} + \int_{-\infty}^{\infty} dJ\rho(J)\Big[\Theta(J-T\epsilon)+\Theta(-J-T\epsilon)\Big]\Big[f(J,h,s)-1\Big], \tag{3.25}$$

$$1<\alpha<2:$$

$$\int_{-\infty}^{\infty} dJ\rho(J)\Big[f(J,h,s)-f'(0,h,s)J-1\Big] = -s^2 C_\alpha \tanh^2(\beta h)\frac{(T\epsilon)^{2-\alpha}}{2-\alpha} + \int_{-\infty}^{\infty} dJ\rho(J)\Big[\Theta(J-T\epsilon)+\Theta(-J-T\epsilon)\Big] \times\Big[f(J,h,s)-f'(0,h,s)J-1\Big]. \tag{3.26}$$

The symbol $\Theta(J)$ denotes the Heaviside step function: $\Theta(J)=1$ when $J>0$ and $\Theta(J)=0$ otherwise. We define the normalized distribution $P_\epsilon(J)$ in terms of $\rho(J)$

$$P_\epsilon(J) \equiv \frac{\alpha(T\epsilon)^\alpha}{2C_\alpha}\rho(J)\Big[\Theta(J-T\epsilon)+\Theta(-J-T\epsilon)\Big]. \tag{3.27}$$

Subsituting equations (3.25) and (3.26) in, respectively, equations (3.22) and (3.23) the integrals over s can be calculated analytically:

$$W_\epsilon(h) = \sum_{k=0}^{\infty}\exp(-c)\frac{c^k}{k!}\int\left(\prod_{r=1}^{k}dh_r W_\epsilon(h_r)\right)\int\left(\prod_{r=1}^{k}dJ_r P_\epsilon(J_r)\right)\int\mathcal{D}z \times\delta\left(h-\tilde{J}_0 m-\beta^{-1}\sum_{r=1}^{k}\operatorname{atanh}\Big[\tanh(\beta J_r)\tanh(\beta h_r)\Big]-\sqrt{2q\Delta}z\right), \tag{3.28}$$

where $\mathcal{D}z = (2\pi)^{-\frac{1}{2}} \exp(-z^2/2)dz$ and:

$$c = \frac{2C_\alpha}{\alpha(T\epsilon)^\alpha}, \qquad \Delta = \frac{(T\epsilon)^{2-\alpha}C_\alpha}{2-\alpha}, \tag{3.29}$$

$$\tilde{J}_0 = \left(J_0 + 2\gamma C_\alpha \left[\frac{(T\epsilon)^{1-\alpha}}{1-\alpha}\right]\right). \tag{3.30}$$

To describe the thermodynamic behavior of Lévy spin glasses one has to solve the set of equations (3.24) and (3.28) for $\epsilon \to 0$. When we compare equation (3.28) with the order parameter equations of FiC systems (2.35), it describes the effective field distribution of a FiC system of Ising spins, in which the number of connections per site k follows a Poissonian distribution with connectivity c. The values of the k couplings attached to a site are drawn from the distribution $P_\epsilon(J)$, see equation (3.27). In addition, the couplings that satisfy $|J| < T\epsilon$ yield an interaction with the global magnetization with effective strength $\tilde{J}_0$ and an extra source of noise represented by the Gaussian random variable z with zero mean and variance Δ, see equation (3.28). The effective strength contains the shift parameter J_0 and a term linear in γ corresponding with the center of the distribution of the couplings.

The interpretation of equation (3.28) is clear for any $\epsilon > 0$: the effective field contains a Poissonian term coming from a finite number of strong bonds and a Gaussian term coming from an infinite number of weak bonds it interacts with. One can perform the limit $\alpha \to 2$ to find the effective field $J_0 m + J_1\sqrt{q}z$, i.e. the RS solution of the SK model, see subsection (2.4.1). Indeed, one has $\Delta \to \frac{J_1^2}{2}$ and $c \to 0$ for $\alpha \to 2$. The equations (3.24) and (3.28) show explicitly how Lévy spin glasses are a hybrid between FC and FiC models. Eq. (3.28) is formally similar to the equation describing the behaviour of composite models [144, 145], where each spin interacts through a finite number of strong couplings and an infinite number of weak couplings. The specific choice of the distribution of coupling strengths in [144, 145] allows an interpolation between the SK model and the Viana-Bray model. When one takes a Gaussian ansatz for the distribution W_ϵ, equation (3.28) becomes equal to the result derived by Cizeau and Bouchaud [28] in the limit $\epsilon \to 0$. The effective connectivity c is plotted as a function of α in figure 3.3. We remark that it is a nonmonotonous function of the variable α. The variance Δ is monotonous as a function of α interpolating between 0 and $J_1^2/2$. The effective connectivity c has the following behavior for $J_0 = 0$ and $\gamma > 0$:

- $\alpha < 1$: it increases monotonously to $\alpha = 1$. For $\alpha \to 1$ it diverges to $+\infty$
- $\alpha > 1$: it is $-\infty$
- $\alpha = 1$: we do not know the precise value

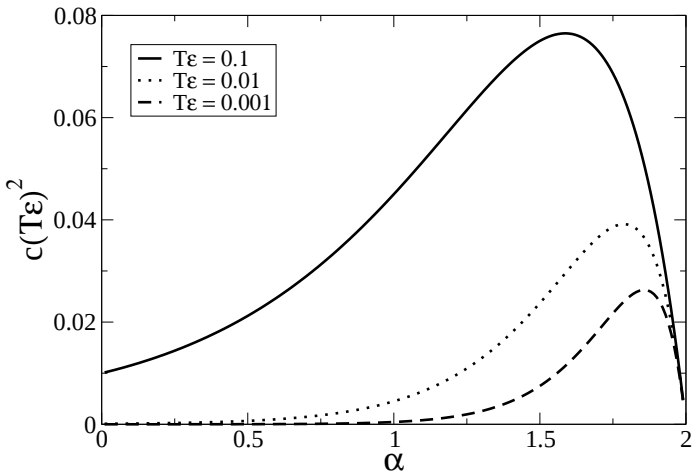


Figure 3.3: The mean connectivity c of the Poissonian backbone graph of strong bonds as a function of α for different values of the cutoff $T\epsilon$. The connectivity reaches a maximum for a value $\alpha^* \in [1.5, 1.8]$

The population dynamics algorithm, see section 2.2.3, can be easily adapted to solve numerically equation (3.28) and obtain $W_\epsilon(h)$. The idea is to obtain numerical results for sufficiently small values of $T\epsilon$ in a way that they can be extrapolated for $T\epsilon \to 0$: the first two moments of the distribution already obtain their limiting values around $T\epsilon \lesssim 0.5$. The equations become very hard to solve around $\alpha \approx 1.5$ because the mean connectivity c has a maximum there, see figure 3.3. For small values of $\alpha \lesssim 0.1$ population dynamics becomes inaccurate because of numerical imprecisions due to the larger tails of the coupling distribution.

In figure 3.4 we plotted the solution to the selfconsistent equation (3.28) for different values of α. The result is compared with the Gaussian ansatz (solid lines), used in [28]. The difference between both approaches is clear. For $\alpha \to 2$ the distribution of fields becomes more and more Gaussian. For $\alpha < 2$ the distributions of fields are not Lévy but leptokurtic distributions where the moments as a function of the size of the population converge to a finite value. Leptokurtic distributions have a smaller kurtosis than a Gaussian distribution with the same variance.

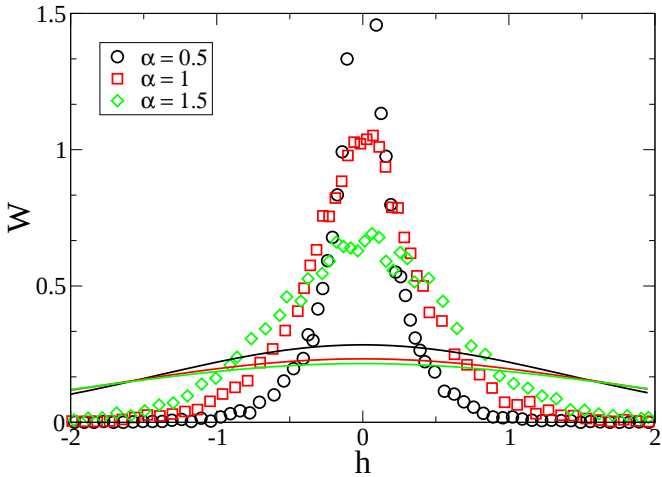


Figure 3.4: The distribution of cavity fields for different values of α with $J_0 = \gamma = 0$, $J_1 = 1$, $T = 0.4$ and several values of α. The markers give the distributions according to equation (3.28) while the solid lines are obtained through the Gaussian assumption made in [28]. All the moments of the distributions are finite. Therefore we have leptokurtic distributions which are neither Gaussian nor Lévy distributions.

3.5 Simulations

The Lévy spin glass has been simulated in [79]. The algorithm used is a variation upon the Metropolis algorithm and makes a distinction between strong and weak bonds, see figure 3.2.

In the algorithm one makes the usual single spin flip updates as done in the Metroplis algorithm and updates of clusters of spins connected through strong bonds. The Monte Carlo algorithm on a lattice of size `N` and for a number of sweeps equal to `SimTime` works as follows:

```
Monte Carlo(N, SimTime, J, Spin)[
        makeCluster(N, J, Clusterlist);
        for(int t=0..SimTime)[
                for(int i=0..N)[
                        with chance p do Metrop(N, J, Spin);
```

```
                        with chance 1-p
                        do Cluster(N, J, Spin, Clusterlist);
                ]
        ]
]
```

The matrix `J` contains the interactions while `Spin` contains the spin configuration.. The algorithm `Metrop(N, J, Spin)` performs a Metropolis single flip move while `Cluster(N, J, Spin)` flips a whole cluster. The cluster is selected from an array of clusterlists `Clusterlist`. The list `Clusterlist[n]` contains all clusters of bonds which have $\beta J > \beta n$. We take $n < K$ with $\beta K > \beta J_*$ the maximum bond strength. This list is made by the algorithm `makeCluster`. The `Cluster` algorithm works as follows:

```
 Cluster(N, J, Spin, Clusterlist)[
         select n from [1, .., K];
         select a cluster C from ClusterList[n];
         Calculate the energy difference E
         when all the spins in cluster are flipped;
         if(E smaller than 0)[
                  accept the move;
         ]
         else[
                   draw a uniform random variable r in [0,1[;
                   if(r smaller than exp(-beta E))[
                             accept the move;
                   ]
         ]
]
```

The cluster moves are important to find the equilibrium configurations. Without these moves the simulation does not reach equilibrium since the clusters of strong bonds do not move. In figure 3.5 we compare the solution of the equations (3.24) and (3.28) with Monte-Carlo simulations. At low temperatures we find a good agreement between the simulations and the theory. Around the critical temperature the magnetization obtained by the simulations is larger than the one predicted by the theory. The reason for this difference is that the simulations equilibrate very slowly. Indeed, as shown in the inset of figure 3.5 the magnetization decays as a power law as a function of the number of Monte-Carlo sweeps. The presence of strong bonds slows down the dynamics since the effect becomes larger for smaller values of α. For very low temperatures the simulation results for the magnetization deviate from those of the RS result. The RS ansatz (3.19) is invalid for very low temperatures, see section 3.8.

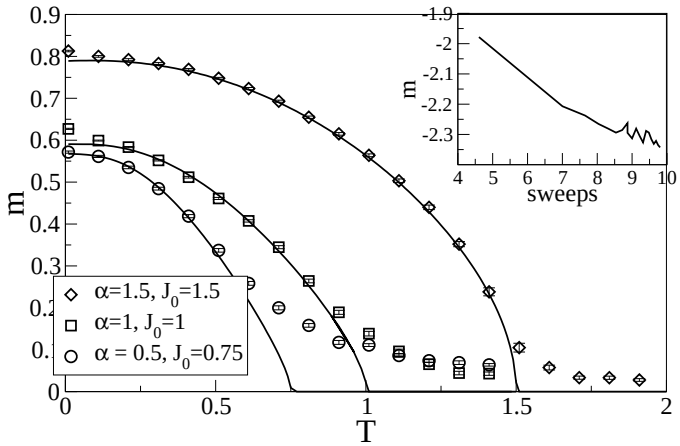


Figure 3.5: The magnetization m as a function of the temperature T for several values of α and J_0. Simulation results (makers) are compared with results from the theory (solid lines) for $J_1 = 1$, $\gamma = 0$. At low temperatures theory and simulations are in good agreement. Because of the increase in the equilibration time around the critical temperature the results from simulation overestimate the magnetization. The inset confirms this: it shows the value of the magnetization as a function of the number of Monte Carlo sweeps for $\alpha = 0.5$, $J_0 = 0.75$ and $T = 1$.

3.6 Cavity method

We have seen that the Gaussian assumption is incorrect. What is wrong in the reasonings made in section 3.3? Why can we not apply the CLT to the infinite sum of random variables in equation 3.11?

Gaussian assumption revisited

We revisit the use of the CLT. The random variable Y

$$Y = N^{-1/\alpha} \sum_{i=1}^{N} Y_i, \tag{3.31}$$

is distributed for $N \to \infty$ according to the Lévy distribution $L_\alpha^{J_1,\gamma,J_0}$, for i.i.d.r.v. Y_i with tails given by equation (3.9).

Let us have a more careful look to equation (3.11). For Gaussian variables, we approximate for $N \to \infty$:

$$h_j = \frac{1}{N^{1/2}} \sum_{i=1}^{N} J_{ij} \tanh(\beta h_i) + \mathcal{O}(1/\sqrt{N}). \tag{3.32}$$

Equation (3.32) is indeed of the form (3.31).

For non-Gaussian variables the assumption $\frac{J_{ij}}{N^{1/\alpha}} \ll 1$, is not anymore valid. Indeed, the probability that $J_{ij} > N^{1/\alpha}$ is of order $\mathcal{O}(1/N)$, see equation (3.6). Therefore, one can not write the equation in the form (3.31). The distribution of the Y_i variables in (3.31) will depend on N.

Modified cavity method

In contrast with [28] we apply the CLT only to the fields coming from the weak bonds, i.e. bonds smaller than a cutoff $T\epsilon$. The bonds larger than $T\epsilon$ form a backbone graph of strong bonds which is treated as a FiC system. For $\epsilon \to 0$ we expect to find the behavior of the spin glass as long as replica symmetry is stable.

The cavity method leads, within one pure phase, to equation (2.87):

$$\begin{aligned} h_j^{(i)} = \beta^{-1} &\sum_{k \in \overline{\omega}^{(i)} \setminus j} \operatorname{atanh}\left(\tanh\left(\beta h_k^{(i)}\right) \tanh\left(\beta \frac{J_{ik}}{N^{1/\alpha}}\right)\right) \\ &+\beta^{-1} \sum_{k \in \omega^{(i)} \setminus j} \operatorname{atanh}\left(\tanh\left(\beta h_k^{(i)}\right) \tanh\left(\beta \frac{J_{ik}}{N^{1/\alpha}}\right)\right), \end{aligned} \tag{3.33}$$

with the sets $\overline{\omega}^{(i)}$ and $\omega^{(i)}$ of respectively, weak and strong bonds, defined as

$$\overline{\omega}^{(i)} \equiv \left\{ j \in \mathbb{N} \cap [1, N] | J_{ij} < T\epsilon N^{1/\alpha} \right\}, \tag{3.34}$$

$$\omega^{(i)} \equiv (\mathbb{N} \cap [1, N]) \setminus \overline{\omega}^{(i)}. \tag{3.35}$$

We treat the term coming from the weak and strong bonds separately. The average over the disorder is taken through the following definitions of the distributions of

cavity fields:

$$W_{\rm w}(h) \equiv \overline{\frac{\sum_{i=1}^{N}\sum_{j\in\overline{\omega}_i}\delta\left(h-h_i^{(j)}\right)}{\sum_{i=1}^{N}\sum_{j\in\overline{\omega}_i}}}, \tag{3.36}$$

$$W_{\rm s}(h) \equiv \overline{\frac{\sum_{i=1}^{N}\sum_{j\in\omega_i}\delta\left(h-h_i^{(j)}\right)}{\sum_{i=1}^{N}\sum_{j\in\omega_i}}}, \tag{3.37}$$

$$W_{\rm g}(z) \equiv \overline{\frac{\sum_{i=1}^{N}\delta(z-z_i^{(j)})}{N}}, \tag{3.38}$$

where we used the notation:

$$z_i^{(j)} = \beta^{-1}\sum_{k\in\overline{\omega}_i\setminus j}\operatorname{atanh}\left(\tanh\left(\beta g_k^{(i)}\right)\tanh\left(\beta J_{ki}\right)\right). \tag{3.39}$$

The difference between the distribution $W_{\rm s}$ and the distribution $W_{\rm w}$ is that the former consists of cavity fields $h_i^{(j)}$ propagating along strong bonds J_{ij} while the latter consists of cavity fields propagating along weak bonds. The distribution $W_{\rm g}$ consists of the terms coming from the weka bonds contributing to $h_j^{(i)}$.

Strong bonds

The backbone graphs of strong bonds form the Poissonian ensemble $\mathcal{G}_{\rm p}(N, c_{\rm b})$, see appendix B.2.2. The mean connectivity $c_{\rm b}$ equals:

$$c_{\rm b} = \lim_{N\to\infty}\frac{\sum_{i=1}^{N}|\omega^{(i)}(\epsilon)|}{N} = \int_{T\epsilon}^{\infty}dJ\rho(J) + \int_{-\infty}^{-T\epsilon}dJ\rho(J) = \frac{2C_\alpha}{\alpha(T\epsilon)^\alpha}, \tag{3.40}$$

with $\rho(J)$ the large tail behavior as defined in (3.9). With the definitions (3.36-3.38) we get

$$\begin{aligned} W_{\rm s}(h) &= \sum_{k=0}^{\infty}\frac{p_{\rm poiss}(k;c_{\rm b})k}{c_{\rm b}}\int\left(\prod_{r=1}^{k-1}dh_r W_{\rm s}(h_r)\right)\int\left(\prod_{r=1}^{k-1}dJ_r P_\epsilon(J_r)\right)\int dz W_{\rm g}(z) \\ &\quad\times\delta\left(h - z - \beta^{-1}\sum_{r=1}^{k-1}\operatorname{atanh}\left(\tanh(\beta h_r)\tanh(\beta J_r)\right)\right). \end{aligned} \tag{3.41}$$

Weak bonds

As the size of the set of weak bonds $|\overline{\omega}_i| \sim \mathcal{O}(N)$, we set $z_i^{(j_1)} = z_i^{(j_2)} = z_i$. The same can be done with the cavity fields $h_i^{(j)}$ propagating along weak bonds to get

$$W_{\rm w}(h) = \sum_{k=0}^{\infty} p_{\rm poiss}(k;c) \int \prod_{r=1}^{k} dh_r W_{\rm s}(h_r) \int \prod_{r=1}^{k} dJ_r P_\epsilon(J_r) \int dz W_{\rm g}(z) \\ \delta\left(h - z - \beta^{-1} \sum_{r=1}^{k} \operatorname{atanh}\left(\tanh(\beta h_r)\tanh(\beta J_r)\right)\right). \quad (3.42)$$

When $\epsilon \ll 1$ we expand the terms in the weak bonds to get

$$z_i = \frac{1}{N^{1/2}} \sum_{k \in \overline{\omega}_i} \frac{J_{ki}}{N^{1/\alpha - 1/2}} \tanh\left(\beta g_k^{(i)}\right), \quad (3.43)$$

with $J_{ki}/N^{1/\alpha} \sim \mathcal{O}(\epsilon)$. According to the CLT the distribution of the variables z_i is given by a Gaussian:

$$W_g(z) = \frac{1}{\sqrt{4\pi\Delta q}} \exp\left(-\frac{(z - \tilde{J}_0 m)^2}{4\Delta q}\right), \quad (3.44)$$

with mean and variance:

$$\tilde{J}_0 m = \left[\int_{-T\epsilon}^{T\epsilon} dJ \left(\frac{P_\alpha^{J_1,\gamma,J_0}(J)}{\int_{-T\epsilon}^{T\epsilon} P_\alpha^{J_1,\gamma,J_0}(J)}\right) \frac{J}{N^{1/\alpha - 1}}\right] \left[\int dh \tanh(\beta h) W_w(h)\right], \quad (3.45)$$

$$2\Delta q = \left[\int_{-T\epsilon}^{T\epsilon} dJ \left(\frac{P_\alpha^{J_1,\gamma,J_0}(J)}{\int_{-T\epsilon}^{T\epsilon} P_\alpha^{J_1,\gamma,J_0}(J)}\right) \frac{J^2}{N^{2/\alpha}}\right] \left[\int dh \tanh^2(\beta h) W_w(h)\right]. \quad (3.46)$$

The integrals over the couplings in equations (3.45) and (3.46) can be calculated using the definitions of the Lévy distributions (3.7, 3.8)., see [127]:

$$\int_{-T\epsilon}^{T\epsilon} dJ P_\alpha^{J_1,\gamma,J_0}(J) \frac{J}{N^{1/\alpha - 1}} = \tilde{J}_0, \quad (3.47)$$

$$\int_{-T\epsilon}^{T\epsilon} dJ P_\alpha^{J_1,\gamma,J_0}(J) \left(\frac{J}{N^{1/\alpha}}\right)^2 = 2\Delta. \quad (3.48)$$

Summary

We use the following property of Poissonian distributions: $\frac{1}{c}p_{\rm poiss}(k;c)k = p_{\rm poiss}(k-1;c)$ to find $W_{\rm w}(g) = W_{\rm s}(g) = W_\epsilon(g)$, i.e. the solutions to (3.41) and (3.42) are the same as the solution W_ϵ of (3.28). Indeed, equations (3.42) combined with (3.45) and (3.46) are identical to equations (3.24) and (3.28) derived with the replica method.

From the cavity approach the importance of the CLT in Lévy spin glasses becomes clear: the couplings have a divergent variance, therefore one can not apply the CLT as done in [28]. We remark that the effective coupling $\tilde{J}_0$ and the parameter 2Δ appearing in the replica method are the mean and the variance of the weak couplings. The distribution $W_\epsilon(h)$ in equations (3.28) is the distribution of the cavity fields propagating along the backbone graph of strong bonds. With the cavity method it becomes straightforward how to generalize this approach to any coupling distribution with tails of the form (3.9). One finds (3.28) but the values of $\tilde{J}_0$ and Δ will differ from the expressions (3.30). One has to calculate (3.47) and (3.48) for the specific distributions of the couplings.

3.7 Stability of the replica symmetric ansatz

We determine the regions of replica symmetry breaking with the two-replica approach, see subsection 2.2.4. The effective field distribution of two sets of uncoupled spins on the same graph with the same set of couplings is given by, see also equation (2.51):

$$\begin{aligned} W_\epsilon(h^1,h^2) = &\int dz^1 dz^2 W_{\rm g}(z^1,z^2) \sum_{k=0}^{\infty} \frac{p_{\rm poiss}(k;c)k}{c} \\ &\times \int \left(\prod_{r=1}^{k-1} dh_r W_\epsilon(h_r^1,h_r^2) \right) \int \left(\prod_{r=1}^{k-1} dJ_r P_\epsilon(J_r) \right) \\ &\times \delta \left(h^1 - z^1 - \beta^{-1} \sum_{r=1}^{k-1} \operatorname{atanh}\left(\tanh(\beta h_r^1) \tanh(\beta J_r) \right) \right) \\ &\times \delta \left(h^2 - z^2 - \beta^{-1} \sum_{r=1}^{k-1} \operatorname{atanh}\left(\tanh(\beta h_r^2) \tanh(\beta J_r) \right) \right) . \end{aligned} \tag{3.49}$$

We assume, just like we did in section 3.6, that we can apply the CLT on the z-fields:

$$W_{\rm g}(z^1,z^2)$$
$$= \frac{1}{4\Delta\pi\sqrt{q^1q^2(1-\rho^2)}}\exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{(z^1-\tilde{J}_0m^1)^2}{2\Delta q^1}+\frac{(z^2-\tilde{J}_0m^2)^2}{2\Delta q^2}\right)\right)$$
$$\times\exp\left(\frac{\rho(z^1-\tilde{J}_0m^1)(z^2-\tilde{J}_0m^2)}{2(1-\rho^2)\Delta\sqrt{q^1q^2}}\right). \tag{3.50}$$

The order parameters become

$$m^1 = \int dg^1dg^2\tanh(\beta g^1)W_\epsilon(g^1,g^2), \tag{3.51}$$
$$q^1 = \int dg^1dg^2\tanh^2(\beta g^1)W_\epsilon(g^1,g^2), \tag{3.52}$$
$$m^2 = \int dg^1dg^2\tanh(\beta g^2)W_\epsilon(g^1,g^2), \tag{3.53}$$
$$q^2 = \int dg^1dg^2\tanh^2(\beta g^2)W_\epsilon(g^1,g^2), \tag{3.54}$$
$$\rho\sqrt{q^1q^2} = \int dg^1dg^2\tanh(\beta g^2)\tanh(\beta g^1)W_\epsilon(g^1,g^2). \tag{3.55}$$

In the limit $\alpha\to 2$ we find $W_\epsilon(h^1,h^2)=W_{\rm g}(h^1,h^2)$. An expansion around the RS solution $m^1=m^2=m$, $q^1=q^2=q$ and $1-|\rho|\sim\mathcal{O}(\delta)$, with $\delta\ll 1$, leads to the following stability condition:

$$\beta^{-2}=\int_{-\infty}^{+\infty}\frac{du}{\sqrt{2\pi}}\exp\left(-\frac{u^2}{2}\right)\mathrm{sech}^4\left(\beta\sqrt{q}u+\beta\tilde{J}_0m\right). \tag{3.56}$$

The parameters (q,m) in (3.56) are, respectively, the overlap parameter q and the magnetization m of the SK model, i.e. the solutions to the equations (2.129) and (2.130). Equation (3.56) is precisely the AT line of the SK model equation (3.56). We discuss in the next section the solution of equations (3.49)-(3.55).

3.8 Phase diagram

Four phases emerge which depend on the values of the order parameters m, q and ρ defined in (3.24) and (3.55): a paramagnetic phase (P) with $m = q = 0$, a ferromagnetic phase (F) with $m > 0, q > 0$ and $\rho = 1$, a mixed phase (M) with $m > 0, q > 0$ and $\rho < 1$ and a spin-glass phase (SG) with $m = 0, q > 0$ and with $\rho < 1$. We did not find a RS SG phase with $m = 0, q > 0$ and $\rho = 1$ or the restoration of the replica symmetry when $T = 0$, contrary to the conjectures made in [28]. The replica symmetry is always broken in the SG phase leading to a phase with many pure states, see figure 3.1.

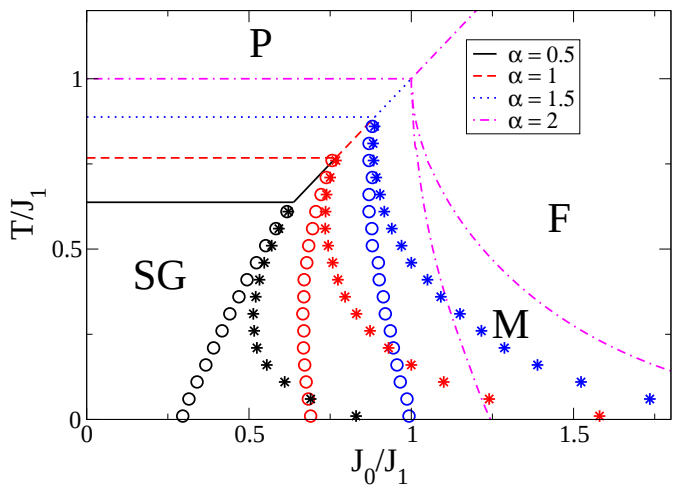


Figure 3.6: The $(T/J_1, J_0/J_1)$ phase diagram of a Lévy spin glass for several values of α and with a skewness $\gamma = 0$. Four phases appear: P (paramagnetic), F (ferromagnetic), M (mixed) and SG (spin glass). The circles present the SG-M transitions (left SG and right M) and the stars indicate were the F phase becomes stable against replica symmetry breaking (left M and right F). For $\alpha = 2$ the phase diagram coincides with that of the SK model.

The ferromagnetic phase contains a region stable to RSB effects (F) and a region unstable to RSB effects which is usually referred to as the mixed phase M.

The P-F and P-SG transitions are determined using an expansion of the self-consistent equation (3.28) around the paramagnetic solution $W_\epsilon(h) = \delta(h)$. For

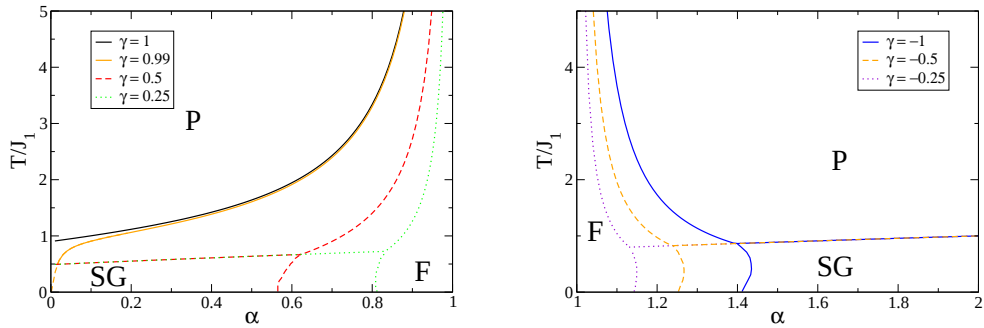


Figure 3.7: The $(T/J_1, \alpha)$ phase diagrams for different values of the skewness γ and a shift $J_0 = 0$. The figure on the left gives the phase diagram for $\gamma > 0$ while the figure on the right gives the phase diagram for $\gamma < 0$. For $\gamma = 1$ and $\alpha < 1$ there is no SG phase. The P-SG transition is independent of γ.

$\gamma = 0$ we find the same bifurcation lines as derived in [79]. To determine the SG-F transition and the M-F transition one has to solve numerically, respectively, equations (3.28) and (3.49) with for instance a population dynamics algorithm.

In figure 3.6 the different phases in the $(J_0/J_1, T/J_1)$ phase diagram are presented for a skewness $\gamma = 0$ and several values of α. The open circles present the SG-F transitions and the stars mark the points where the F phase becomes stable with respect to RSB. These results generalize the phase diagram obtained in the paper of Sherrington and Kirkpatrick [161] to coupling distributions with a large tail. The P-F transition is independent of α while the SG phase increases in favor of a smaller F phase as a function of α. The RSB effects decrease when α decreases: indeed the M-phase becomes smaller and the reentrance effect in the SG-F phase transition line diminishes and finally disappears. This is related to the decrease of frustration due to the presence of stronger bonds that dominate the systems behavior. The decrease of RSB is similar to the decrease of RSB when the connectivity of FiC systems decreases, see figure 2.14. All phase transitions are continuous.

In figure 3.7 we present the $(T/J_1, \alpha)$-phase diagram for different values of γ and $J_0 = 0$. We consider the following regions:

- $\gamma > 0$ and $\alpha < 1$ (left figure): the F phase increases and the SG phase decreases as a function of increasing γ. The SG phase disappears at $\gamma = 1$. For values of γ very close to one the SG phase is only present for very small values of α. The transition temperature between the P and F phase becomes

infinite for $\alpha \to 1^-$.

- $\gamma < 0$ and $\alpha > 1$ (right figure): the F phase decreases and the SG phase increases as a function of increasing γ. The transition temperature between the P and F phase becomes infinite for $\alpha \to 1^+$.
- $\gamma > 0$ and $\alpha > 1$ (not shown): there is no F phase but a P and SG phase.
- $\gamma < 0$ and $\alpha < 1$ (not shown): there is no F phase but a P and SG phase.

We have some additional remarks: The transition temperature becomes very large for $\alpha \to 1^{\pm}$ (for, respectively, $\gamma < 0$ and $\gamma > 0$) because the effective coupling $\tilde{J}_0 \to \infty$. There is no SG phase for $\gamma = 1$ and $\alpha < 1$ because there are no negative couplings, only the P-F transition occurs. The P-SG transitions coincide for different values of γ. For small values of α the results of population dynamics become inaccurate because of numerical imprecisions when dealing with a broad range of coupling values. In this case we have used the instability line of the P phase with respect to the F phase as of the location of the SG-F transition.

3.9 The free energy and the entropy

It is possible to calculate the free energy from the saddle point equations. We use the RS ansatz and we introduce again a cutoff ϵ. The RS free energy equals:

$$f_{\rm RS} = \lim_{\epsilon \to 0} \left(\frac{J_0}{2} m^2 - \frac{\beta}{2} \Delta (1-q)^2 + f_{\rm site}(\epsilon) - \frac{c}{2} f_{\rm link}(\epsilon) \right), \tag{3.57}$$

with

$$\begin{aligned} f_{\rm site}(\epsilon) \;=\; & -\beta^{-1} \sum_{k=0}^{\infty} \frac{c^k e^{-c}}{k!} \prod_{l=1}^{k} \left[\int dh_l W(h_l) \int d\hat{r}_l P_\epsilon(\hat{r}_l) \right] \\ & \times \int Dz \log \left(\sum_{\sigma} \exp \left((\beta J_0 m + \sqrt{2q\Delta} z) \sigma \right) \right. \\ & \left. \times \prod_{\ell=1}^{k} \left(\sum_{\tau} \exp\left(\beta \hat{r}_\ell \tau \sigma\right) \exp\left(\beta h_\ell \tau\right) \right) \right), \end{aligned} \tag{3.58}$$

and

$$f_{\rm link}(\epsilon) = -\beta^{-1} \int dh dh' W(h) W(h') \int d\hat{r} P_\epsilon(\hat{r}) \ln \Big[\sum_{\sigma,\tau} \exp\left(\beta \hat{r} \sigma\tau\right) \exp\left(\beta h \sigma\right) \exp\left(\beta h' \tau\right) \Big]. \tag{3.59}$$

The entropy is given by $s = \beta^2 \frac{\partial f}{\partial \beta} = \lim_{\epsilon \to 0} s(\epsilon)$ with $s(\epsilon)$:

$$s(\epsilon) = \beta^2 \frac{\Delta}{2} \left(1 - q^2\right) - \beta^2 \Delta \left(1 - q\right) + s_{\rm site}(\epsilon) - \frac{c}{2} s_{\rm link}(\epsilon). \tag{3.60}$$

The quantity $s_{\rm link}$ is equal to

$$s_{\rm link}(\epsilon) = - \int dh dh' W_\epsilon(h) W_\epsilon(h') \int dJ P_\epsilon(J) \sum_{\sigma,\tau} \frac{\exp\left(\beta J \sigma\tau + \beta h \sigma + \beta h' \tau\right)}{\sum_{\sigma,\tau} \exp\left(\beta J \sigma\tau + \beta h \sigma + \beta h' \tau\right)} \log \left(\frac{\exp\left(\beta J \sigma\tau + \beta h \sigma + \beta h' \tau\right)}{\sum_{\sigma,\tau} \exp\left(\beta J \sigma\tau + \beta h \sigma + \beta h' \tau\right)} \right), \tag{3.61}$$

and $s_{\rm site}$ reads

$$s_{\rm site}(\epsilon) = - \sum_{k=0}^{\infty} \frac{c^k e^{-c}}{k!} \prod_{l=1}^{k} \Big[\int dh_l W_\epsilon(h_l) \int dJ_l P_\epsilon(J_l) \Big] \int \mathcal{D}z \sum_{\sigma;(\tau_1,\tau_2,\cdots,\tau_k)} \frac{\exp\Big((\beta J_0 m + \sqrt{2q\Delta} z) \sigma \Big) \prod_{\ell=1}^{k} \left(\exp\left(\beta J_\ell \tau_\ell \sigma\right) \exp\left(\beta h_\ell \tau_\ell\right) \right)}{\sum_{\sigma;(\tau_1,\tau_2,\cdots,\tau_k)} \exp\Big((\beta J_0 m + \sqrt{2q\Delta} z) \sigma \Big) \prod_{\ell=1}^{k} \left(\exp\left(\beta J_\ell \tau_\ell \sigma\right) \exp\left(\beta h_\ell \tau_\ell\right) \right)} \log \left[\frac{\exp\Big((\beta J_0 m + \sqrt{2q\Delta} z) \sigma \Big) \prod_{\ell=1}^{k} \left(\exp\left(\beta J_\ell \tau_\ell \sigma\right) \exp\left(\beta h_\ell \tau_\ell\right) \right)}{\sum_{\sigma;(\tau_1,\tau_2,\cdots,\tau_k)} \exp\Big((\beta J_0 m + \sqrt{2q\Delta} z) \sigma \Big) \prod_{\ell=1}^{k} \left(\exp\left(\beta J_\ell \tau_\ell \sigma\right) \exp\left(\beta h_\ell \tau_\ell\right) \right)} \right] \tag{3.62}$$

For $\alpha \to 2$ one gets precisely the expressions for the free energy and the entropy of the SK model [83] (see also subsection 2.4.1). The free energies $f_{\rm site}$ ($= \langle \Delta F_{\rm site} \rangle$) and $f_{\rm link}$ ($= \langle \Delta F_{\rm link} \rangle$) and the entropies $s_{\rm site}$ ($= \langle \Delta S_{\rm site} \rangle$) and $s_{\rm link}$ ($= \langle \Delta S_{\rm link} \rangle$) correspond with the ensemble averaged free energy and entropy differences when performing, respectively, a site addition and a link addition on the backbone graph of strong bonds, see subsection 2.3.1. Similar to the selfconsistent equation (3.28)

we find the entropy of an Ising model on a Poissonian graph with mean connectivity c, a distribution of the bonds P_ϵ and an extra Gaussian noise z.

We plotted the entropy s as a function of α in figure 3.8. From this figure we see that the entropy gets less negative, for $T \to 0$ and for smaller values of α and becomes eventually zero for $T \to 0$ and values $\alpha \lesssim 1$. This is consistent with a decrease of RSB effects when α decreases.

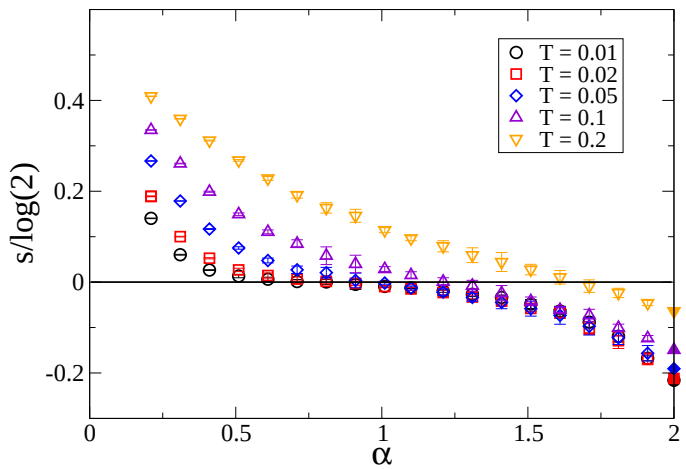


Figure 3.8: The entropy s of a Lévy spin glass as a function of the exponent α for different values of the temperature T, $J_0 = 0$, $\gamma = 0$ and $J_1 = 1$. The filled markers at $\alpha = 2$ show the SK values. The entropy converges to the SK value for $\alpha \to 2$.

3.10 Conclusion

In this chapter we have shown how to derive the phase diagrams of Lévy spin glasses where the couplings between the spins are drawn from a distribution with power law tails characterized by an exponent α. These models are known to have a finite number of strong bonds of order $\mathcal{O}(1)$ and an infinitely amount of weak bonds of order $\mathcal{O}(N^{-1/\alpha})$. This fact is the main ingredient of our solution. Indeed, an explicity distinction between strong and weak bonds is made in our solution through the replica method, Monte Carlo method or the cavity method.

The crucial difference with previous works [28] and [79] is that we derive the phase diagrams, the entropy and the stability against replica symmetry of Lévy spin glasses without using the Gaussian assumption for the distribution of fields. We have neither found evidence for a replica symmetric spin glass phase, nor for a restoration of the replica symmetry at zero temperature, as was conjectured in [28]. We have used the replica and the cavity method within, respectively, the replica symmetric assumption and the assumption of one pure phase. The resultant effective equations for the distribution of cavity fields show clearly the hybrid character of the model being a mixture between a finite connectivity model and a FC model.

The phase transitions are qualitatively similar to the ones found in the SK model. Large tails do influence quantitatively the phase diagram: the Lévy spin glass model becomes more stable with respect to replica symmetry breaking and the SG phase decreases when the tails get larger. Moreover, the reentrance effects in the replica symmetric phase diagram disappear for $\alpha \lesssim 1$. The replica symmetry breaking transitions are all continuous. The skewness γ in the Lévy distribution can have a big influence on the size of the F phase. For $\alpha \to 2$ the effective distribution of fields becomes Gaussian and we have found back the results of the SK model. For $\alpha < 2$ this distribution is neither Lévy nor Gaussian, but a distribution with finite moments and a kurtosis smaller than a Gaussian with the same variance.

3.11 Some future prospects

Now that one understands how to solve mean-field models with interactions drawn from Lévy distributions one can think of many interesting related questions, for instance the study of the dynamics of these models and the study of the related p-spin model for $p \to \infty$

3.11.1 Dynamics of Lévy processes

As already seen above, Lévy spin glasses have a slow dynamics, see figure 3.5. It would be interesting to investigate this point more quantitively and compare with the SK model. We can use the generating functional analysis like in [74] or the cavity method combining the methods as layed out in [126] and [127].

In the former one calculates:

$$Z(\boldsymbol{\psi}) = \overline{\exp\left(i \sum_{s=1}^{t} \sum_{i=1}^{N} \psi_i \sigma_i(s)\right)}. \tag{3.63}$$

We remark that in the expression (3.63) the i is there from the beginning, hence we do not need an analytic continuation to real numbers. We find the following expression in the case of parallel Glauber dynamics:

$$
\begin{aligned}
P_{0..t}(\sigma^{0..t}|\theta^{0..t-1}) &= p(\sigma(0)) \\
&\times \int \frac{dh^{0..t} d\hat{h}^{0..t}}{2\pi} \prod_{s=0}^{t} \frac{\exp\left(i\hat{h}(s)(h(s)-\theta(s)) + \beta\sigma(s+1)h(s)\right)}{2\cosh\left(\beta h(s)\right)} \\
&\times \exp\left[-iJ_0 \sum_{s=0}^{t} \hat{h}(s) m(s)\right] \\
&\times \exp\left[C(\alpha) \int \frac{dJ}{|J|^{\alpha+1}}\right. \\
&\left.\times \left(\sum_{\tau^{0..t-1}} P_{0..t-1}(\tau^{0..t-1}|J\sigma^{0..t-2}) \exp\left(-iJ\sum_{s=0}^{t} \hat{h}(s)\tau(s)\right) - 1\right)\right] \qquad (3.64)
\end{aligned}
$$

Equation (3.64) can be compared with the static equations (3.22) and (3.23). We encounter again the same problem: the integration over the couplings is not properly normalized. Again one can introduce a cutoff $T\epsilon$ in the distribution of couplings to find:

$$
\begin{aligned}
P_{0..t}(\sigma^{0..t}|\theta^{0..t-1}) &= p_0(\sigma(0)) \sum_{r=0}^{\infty} \exp(-c) \frac{c^r}{r!} \int \prod_{\ell=1}^{r} \left(dJ_\ell P_\epsilon(J_\ell)\right) \\
&\times \sum_{\tau_1^{0..t}, \cdots, \tau_r^{0..t}} \prod_{r=1}^{\ell} P_{0..t-1}\left(\tau_r^{0..t-1}|J_r\sigma^{0..t-2}\right) \\
&\times (\det \boldsymbol{C})^{1/2} \left(\frac{1}{4\pi\Delta}\right)^{t/2} \int \left(\prod_{s=0}^{t-1} dx(s)\right) \exp\left(-\frac{1}{4\Delta}\sum_{s,s'} x(s) C^{-1}(s,s') x(s')\right) \\
&\times \prod_{s=1}^{t} \frac{\exp\left[\beta\sigma(s)\left(J_0 m(s-1) + \sum_{\ell=1}^{r} J_\ell \tau_\ell(s) + x(s) + 2\Delta \sum_{s'<s} R(s,s')\right)\right]}{2\cosh\left[\beta\left(J_0 m(s-1) + \sum_{\ell=1}^{r} J_\ell \tau_\ell(s) + x(s) + 2\Delta \sum_{s'<s} R(s,s')\right)\right]}
\end{aligned}
\qquad (3.65)
$$

with $P_\epsilon(J)$ as defined in (3.27) and

$$m(s) = \sum_{\sigma^{0..s}} \sigma(s) P_{0..s}\left(\sigma^{0..s}|0^{0..s}\right), \tag{3.66}$$

$$C(s,s') = \sum_{\sigma^{0..\max(s,s')}} \sigma(s)\sigma(s') P_{0..\max(s,s')}\left(\sigma^{0..s}|0^{0..s}\right), \tag{3.67}$$

$$R(s,s') = \Theta(s-s')\frac{\partial}{\partial\theta(s')} \sum_{\sigma^{0..s}} \sigma(s) P_{0..s}\left(\sigma^{0..s}|0^{0..s}\right). \tag{3.68}$$

Analogously to the static case, we see that the equations describing the dynamics of Lévy spin glasses, are similar to the equations describing the dynamics of a spin model on a FiC backbone of strong bonds with some additional FC Gaussian noise. For $\alpha \to 2$ we find back the equations describing the parallel Glauber dynamics of the SK model.

3.11.2 p-spin Lévy spin glasses

The spin glass phase and the phase transition to the SG phase are different in models with p-spin interactions than the corresponding behaviour in models with two-spin interactions, see subsections 2.4.1 and 2.4.2. For Gaussian coupling distributions one retrieves the random energy model for $p \to \infty$. How does the SG phase changes for p-spin models when one uses Lévy coupling distributions and what is the corresponding model when $p \to \infty$?

The latter question is not trivial since the Lévy model does not converge to a random energy model for $p \to \infty$. The p-spin model is defined through the Hamiltonian

$$H\left(\{\sigma_i\}_{i=1..N}\right) = -\frac{(p!)^{1/\alpha}}{N^{(p-1)/\alpha}} \sum_{i_1<i_2<\cdots<i_p} J_{i_1 i_2 \cdots i_p}\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_p}. \tag{3.69}$$

with the couplings $J_{i_1 i_2 \cdots i_p}$ distributed according to $P_\alpha^{J_1,0,0}$. We consider couplings drawn from Lévy distributions with $J_0 = \gamma = 0$. The energy density $\rho(E)$ is given by

$$\begin{aligned}\rho(E) &\equiv 2^{-N}\overline{\sum_{\{\sigma_i\}_{i=1..N}} \delta\left(E - H\left(\{\sigma_i\}_{i=1..N}\right)\right)} \\ &= P_\alpha^{J_1,0,0}\left(N^{-1/\alpha}E\right).\end{aligned} \tag{3.70}$$

The two-point energy density is given by

$$\begin{aligned}\rho(E_1,E_2) &= 2^{-2N}\overline{\sum_{\{\sigma_i\}_{i=1..N},\{\tau_i\}_{i=1..N}} \delta\left(E_1 - H\left(\{\sigma_i\}_{i=1..N}\right)\right)\delta\left(E_2 - H\left(\{\tau_i\}\right)_{i=1..N}\right)} \\ &= 2P_\alpha^{J_1,0,0}\left(\frac{E_1+E_2}{2N^{1/\alpha}}\right)P_\alpha^{J_1,0,0}\left(\frac{E_1-E_2}{2N^{1/\alpha}}\right). \qquad (3.71)\end{aligned}$$

For $\alpha = 2$ we get indeed the random energy model, but for other values of α the energies are correlated. Is it possible to take the limit $p \to \infty$ explicitly and identify the model we get then?

3.11.3 Rigorous inequalities

Using the interpolating method one can prove for the SK-model [71, 70], the Viana-Bray model [55] but also for Ising models on a given degree distribution [57] the following properties:

- the limit $\lim_{N\to\infty} \overline{f}$ exists and f is self-averaging.
- the disordered averaged free energy $\overline{f}$ is lower bounded by the RS expression of the free energy $f_{\rm RS}$.

In this chapter we have introduced a cutoff $T\epsilon$ in the distribution of the couplings to decompose the Lévy spin glass in a FC Gaussian background and a FiC backbone of strong couplings, see figure 3.2. It would be interesting to use this decomposition within the interpolating method and combine results obtained in [70] and [55] . Is it possible to prove that the RS expression (3.57) is a lower bound to the disordered free energy $\overline{f}$?

Chapter 4

Parallel dynamics of spin models on graphs

In this chapter we study the parallel dynamics of spin models on a given graph instance. Previous works on parallel dynamics of finitely-connected systems have been concentrated on ensemble-averaged equations [74, 111].

Our aim is to find a message-passing algorithm that determines the marginals of the steady-state distribution of non-equilibrium models. Such an algorithm generalizes the belief-propagation algorithm for equilibrium models and would be an important tool in the study of neural networks of small organsims or the study of congestion problems in traffic networks.

We start the development of such an algorithm from the study of simple models whose steady state is not known. We study the Glauber dynamics of spin models on graphs which contain asymmetric and symmetric links. The neural network of biological organisms and transportation networks are examples of such graphs [175, 81].

On fully symmetric graphs the dynamics fulfills detailed balance and the equilibrium distribution is known. Asymmetric links break the detailed balance condition. We derive a message-passing algorithm in the marginals of the path distribution. The stationary limit is taken taken approximatively through the assumption that the marginal path distributions factorize in time. This procedure leads to a message-passing algorithm in the marginals of the steady state. The results for the magnetization are compared with simulation results. For fully symmetric graphs and fully asymmetric graphs the resultant equations reduce to the corresponding belief-propagation equations. Therefore, the algorithm is exact in those limiting cases.

As an application we discuss the influence of the architecture of the graph on the retrieval properties of associate memories.

4.1 Aim and motivation

In this section we study the dynamics of spin models on graphs. This leads to a couple of interesting questions which we discuss below.

4.1.1 Dynamics on FiC

There exist no efficient tools to study the dynamics of spin models on graphs, contrary to the statics as explained in chapter 2. Several attempts have been made to solve the dynamics of FiC with partial successes. Most of these studies concentrate on ensemble-averaged equations. Moreover, the resultant equations are very complicated making these methods less useful in practice. We give a short overview.

- **Generating functional methods**: The first class of methods are called the generating functional methods (GFA). One calculates the Martin-Siggia-Rose functional [99]. For binary spins with parallel Glauber dynamics this calculation has been done in [74] and for continuous spin models with Langevin dynamics in [116]. The resultant equations are in general too complicated to make useful theoretical predictions. Getting results from simulations is in general much easier.

  In [74] an iterative equation has been derived for the ensemble-averaged single site probability $\overline{P}\left(\sigma^{0..t}|\theta^{1..t}\right)$ of a path $\sigma^{0..t}$ in an external field $\theta^{0..t}$

$$\begin{aligned}\overline{P}\left(\sigma^{0..t}|\theta^{1..t}\right) = \sum_{k\geq 0}\frac{e^{-c}c^k}{k!}\int\left(\prod_{\ell=1}^{k}dJ_\ell P_I(J_\ell)\right)\\ \times\prod_{\ell=1}^{k}\sum_{\sigma_\ell^{0..t-1}}\left[\epsilon\overline{P}\left(\sigma_\ell^{0..t-1}|J_\ell\sigma^{0..t-2}\right)+(1-\epsilon)\overline{P}(\sigma_\ell^{0..t-1}|0^{0..t-1})\right]\\ \times p(\sigma(0))\prod_{s\geq 0}^{t-1}\frac{\exp\left[\beta\sigma(s+1)\left(\theta(s+1)+\sum_{0<\ell'\leq k}J_{\ell'}\sigma_{\ell'}(s)\right)\right]}{2\cosh\left[\beta\left(\theta(s+1)+\sum_{0<\ell'\leq k}J_{\ell'}\sigma_{\ell'}(s)\right)\right]}\end{aligned} \tag{4.1}$$

  Equation (4.1) holds for parallel Glauber dynamics on graphs drawn from the ensemble $\mathcal{G}_{\mathrm{p}}\left(N,c,\epsilon\right)$ defined in appendix B.2.4. The ensemble averaged

magnetization at time step t, $\overline{m}(t)$ is given by

$$m(t) = \sum_{\sigma^{0..t-1}} \overline{P}\left(\sigma^{0..t}|0^{1..t}\right), \tag{4.2}$$

with $0^{1..t}$ the null vector. As the set of equations (4.1) involves a summation of 2^t terms, it is a very complex equation which is in practice not completely solvable. One can solve equation (4.1) up to 10 time steps.

- **Dynamics of spherical models**: Spherical models are exactly solvable models. The Langevin dynamics of spins can be reduced to the memory dependend dynamics of a single spin [157]. This effective dynamics depends on the eigenvalue spectrum of the interaction matrix of the Hamiltonian. In [157] the dynamics has been solved approximately using the effective medium approximation for the spectrum [156]. However, using the exact results for the spectrum of a FiC matrix [88, 150] it must be possible to calculate the dynamics exactly.
- **Dynamical replica analysis** [73, 122]: In this method one assumes that the path distribution $p_t\left(\{\sigma_i\}_{i=1..N}\right)$ depends only on a finite number of selfaveraging macroscopic observables. Within these assumptions the dynamical evolution of these observables is closed. This method allows one to investigate sequential dynamics of spin models.

In this chapter we study the parallel Glauber dynamics with the cavity method. The resultant equations solve the dynamics of spin models on a given graph instance. When taking the ensemble average over Poissonian graphs we find the resultant equations (4.1) determined through the GFA. With the cavity method the meaning of the dynamical order parameters found with the GFA becomes transparent. Since one finds back the same equations as those found with the GFA, the computational complexity of the resultant equations scales exponential in the discrete time. Therefore, one can only compare the results of these equations with Monte Carlo simulations for a couple of time steps.

4.1.2 Dynamics on asymmetric graphs

The dynamics on fully asymmetric graphs can be solved exactly; using an annealed approach the work of Derrida [39, 40] solves the dynamics on asymmetric graphs. In the annealed dynamics the interactions are drawn every time step independently from a random distribution. In a probabilistic approach the dynamics can be solved easily, since the resultant equations (4.1) become Markovian. In our work we go a step further by providing a MPA that determines the marginals on arbitrary fully asymmetric graphs with a local tree structure. We call this algorithm asymmetric belief propagation (ABP).

4.1.3 Message passing algorithms on partially asymmetric graphs

If the dynamics fulfils detailed balance, the stationary distribution is known and the cavity method for equilibrium models, see section 2.3, can be applied to study the stationary behavior. We are interested in cases where the detailed balance condition is not satisfied. A simple example is the parallel Glauber dynamics defined on a non-symmetric graph. Non-symmetric graphs are ubiquitous in experiments. The transportation network of global cargo ship movements [81], neuronal networks of a simple organism [175], protein-protein interaction networks [51, 80] and social networks like the world wide web [143] are all examples of partially asymmetric graphs.

In [175] the neuronal network of the nematode Caenorhabditis elegans has been completely mapped. It consists of 302 neurons wired by about 5000 chemical synapses, 2000 neuromuscular junctions and 600 gap junctions. The chemical synapses are unidirectional connections while the gap junctions are bidirectional. The neoromuscular junctions connect these neurons with muscle cells.

Protein-protein interaction networks contain connections with both undirected as well as directed edges. Transcriptional regulatory networks are directed while interactions coming from biochemical reactions or formation of proteincomplexes are undirected. Combining these networks one finds a hybrid network with both directional and unidirectional links.

We want to generalize the BP algorithm (2.88) to an algorithm that determines the marginals of the stationary states of non-equilibrium models. We study the most simple non-trivial problem: an Ising model on a non-symmetric graph. We take the stationary limit through an approximative method and find an efficient algorithm that determines the marginals of the stationary state of these models. In the limiting case of fully symmetric and fully asymmetric graphs we find back the BP-algorithm and the ABP-algorithm. Theoretical results are compared with simulations.

4.2 Dynamics on a given graph instance

4.2.1 Some definitions and notations

We consider Ising models defined on a given graph instance $G = (V, E)$, with V and E respectively the set of vertices (or sites) and the set of edges. We limit ourselves to simple directed graphs G determined by a connectivity matrix $\boldsymbol{C}$, with elements $[\boldsymbol{C}]_{ij} = c_{ij} \in \{0, 1\}$. When $c_{ij} = 1$ and $c_{ji} = 0$ the graph has a directed edge from the i-th site to the j-th site. When $c_{ij} = c_{ji} = 1$ there is an undirected edge between i and j and when $c_{ij} = c_{ji} = 0$ there are no edges between them.

We define the sets E, $E^{\rm d}$ and $E^{\rm sym}$ through: $E \equiv \{(i,j) \in V \times V | c_{ij} = 1\}$, $E^{\rm d} \equiv \{(i,j) \in V \times V | c_{ij} = 1, c_{ji} = 0\}$ and $E^{\rm sym} \equiv \{(i,j) \in V \times V | c_{ij} = 1, c_{ji} = 1\}$. We study the evolution of Ising like models of binary variables $\sigma_i(t) \in \{-1,1\}$ or pairs of binary variables $(\sigma_i(t), \tau_i(t)) \in \{-1,1\} \times \{-1,1\}$, with $i = 1, \ldots N$ and t the corresponding discrete time step. The dynamics in discrete time is defined by a transition probability $W\left(\{\sigma(s)\}_{i=1..N} | \{\sigma_i(s-1)\}_{i=1..N}\right)$, from the state $\{\sigma_i(s-1)\}_{i=1..N}$ on the $(s-1)$-th time step to the state $\{\sigma_i(s)\}_{i=1..N}$ on the s-th time step. We consider transition probabilities W of the form:

$$\begin{aligned} & W\left[\{\sigma_i(s)\}_{i=1..N} | \{\sigma_i(s-1)\}_{i=1..N}; \{\theta\}_{i=1..N}\right] \\ &= \prod_{i=1}^{N} W\left[\sigma_i(s) | \{\sigma_i(s-1)\}_{i=1..N}; \theta_i(s)\right] \\ &= \prod_{i=1}^{N} W\left[\sigma_i(s) | h_i(s)\right]. \end{aligned} \tag{4.3}$$

The local field $h_i(s)$ is defined through

$$h_i(s) = \sum_{j \in \partial_i^{\rm in}} h_{j \to i}(\sigma_j(s-1)) + \theta_i(s), \tag{4.4}$$

where the field $h_{j \to i}(\sigma_j(s-1))$ quantifies the influence of the spin on the j-th site on the spin on the i-th site and $\theta_i(s)$ is an external field. At this moment $h_{j \to i}$ is an arbitrary function of $\sigma_j(s-1)$. We have used the notation $\partial_i^{\rm in}$ for the neighborhood of all the vertices that influence i directly, i.e. $\partial_i^{\rm in} \equiv \{j \in V | c_{ji} = 1\}$. We also use: $\partial_i^{\rm out} \equiv \{j \in V | c_{ij} = 1\}$, $\partial_i \equiv \partial_i^{\rm in} \cup \partial_i^{\rm out}$ and $\partial_i^{\rm sym} \equiv \partial_i^{\rm in} \cap \partial_i^{\rm out}$. The probability to have the path $\left\{\sigma_i^{t_0..t}\right\}_{i=1..N}$, with $\sigma_i^{t_0..t} = (\sigma_i(t_0), \cdots, \sigma_i(t))$, from time step t_0 to time step t, is given by

$$\begin{aligned} & P_{t_0..t}\left(\left\{\sigma_i^{t_0..t}\right\}_{i=1..N} | \left\{\theta_i^{t_0+1..t}\right\}_{i=1..N}\right) \\ &= \left(\prod_{s=t_0+1}^{t} W\left[\{\sigma_i(s)\}_{i=1..N} | \{\sigma_i(s-1)\}_{i=1..N}; \{\theta_i(s)\}_{i=1..N}\right] \right) \\ &\qquad \times P_{t_0}\left(\{\sigma_i(t_0)\}_{i=1..N}\right), \end{aligned} \tag{4.5}$$

with $P_{t_0}\left(\{\sigma_i(t_0)\}_{i=1..N}\right)$ the probability distribution of the spins at time step t_0.

4.2.2 Dynamical version of the cavity equations

We use the cavity method, see section 2.3, to solve the parallel dynamics on graphs. We derive the dynamical version of the cavity equation using a Bayesian approach analogous to the one followed in subsection 2.3.1. We write the following relationship between a path probability $P_{0..t}$ on the graph G and the probability $P^{(i)}_{0..t}$ on its related cavity graph $G^{(i)}$

$$\begin{aligned} & P_{0..t}\left(\left\{\sigma_j^{0..t}\right\}_{j=1..N} | \left\{\theta_j^{1..t}\right\}_{j=1..N}\right) \\ &= P^{(i)}_{0..t}\left(\left\{\sigma_j^{0..t}\right\}_{j=1..N} | \left\{\theta_j^{1..t}\right\}_{j=1..N} + \left\{\zeta_j^{(i),1..t}\right\}_{j=1..N}\right) \\ &\quad \times \left(\prod_{s=1}^{t} W\left[\sigma_i(s) | \left\{\sigma_j(s-1)\right\}_{j=1..N}; \theta_i(s)\right]\right) p_0(\sigma_i(0)). \end{aligned} \tag{4.6}$$

In equation (4.6) we have introduced an extra field $\zeta_j^{(i)}(s)$, representing the influence of the i-th spin on its neighbours $j \in \partial_i^{\text{out}}$:

$$\zeta_j^{(i)}(s) = h_{i\to j}(\sigma_i(s-1)). \tag{4.7}$$

We have taken a factorized initial distribution P_0:

$$P_0(\{\sigma_i(0)\}_{i=1..N}) = \prod_{i=1}^{N} p_0(\sigma_i(0)). \tag{4.8}$$

The single site marginal $P_{i,0..t}$ has been obtained by summing $P_{0..t}$ in (4.6) over all paths $\sigma_j^{0..t}$ with $j \neq i$. We use the notations

$$\sigma_S = (\sigma_{i_1}, \sigma_{i_2}, \cdots, \sigma_{i_{|S|}}), \tag{4.9}$$

$$P_S\left(\sigma_S^{0..t} | \left\{\theta_i^{1..t}\right\}_{i=1..N}\right) = \sum_{\sigma_i^{0..t},\, i\neq S} P\left(\left\{\sigma_i^{0..t}\right\}_{i=1..N} | \left\{\theta_i^{1..t}\right\}_{i=1..N}\right), \tag{4.10}$$

with S a set of indices: $S = \left\{i_1, \cdots, i_{|S|}\right\}$, where $|S|$ denotes the size of the set S. Within this notation $P_{\partial_i,0..t}$ is the joint probability of the paths on the neighbours of i. When we sum over all paths $\sigma_j^{0..t}$, with $j \neq i$, on the left and right hand side

of (4.6), we get

$$\begin{aligned} P_{i,0..t}\left(\sigma_i^{0..t}|\left\{\theta_j\right\}_{j=1..N}\right) &= \sum_{\tau_{\partial_i}^{0..t}} P_{\partial_i\cup i,0..t-1}\left(\tau_{\partial_i}^{0..t},\sigma_i^{0..t}\mid\left\{\theta^{1..t}\right\}_{j=1..N}\right) \\ &= \sum_{\tau_{\partial_i}^{0..t}} P^{(i)}_{\partial_i,0..t}\left(\tau_{\partial_i}^{0..t-1}\mid\left\{\theta_j^{1..t-1}+\zeta_j^{(i),1..t-1}\right\}_{j=1..N}\right) \\ &\quad\times\prod_{s=1}^{t} W\left[\sigma_i(s)|h_i(s)\right]p_0(\sigma_i(0)). \end{aligned} \tag{4.11}$$

In the sequel we drop the subscript i in the argument of $P_{i,0..t}$. We make the Bethe-Peierls approximation: i.e. we assume that the spins in the neighborhood ∂_i of i become independent when we remove the i-th spin:

$$P^{(i)}_{\partial_i,t}\left(\tau_{j_1}^{0..t},\cdots,\tau_{j_{|\partial_i|}}^{0..t}|\left\{\theta_j^{1..t}+\zeta_j^{(i),1..t}\right\}_{j=1..N}\right)=\prod_{j\in\partial_i}P^{(i)}_{j,0..t}\left(\tau_j^{0..t}|\theta_j^{1..t}+\zeta_j^{(i),1..t}\right), \tag{4.12}$$

with $\partial_i=\left\{j_1,\cdots,j_{|\partial_i|}\right\}$. The "dynamic Bethe-Peierls condition" Equation (4.12) is valid when the equilibration time is smaller then the loop length. In (4.12) we have taken $\theta_j=0$ when $j\notin\partial i\cup i$. We substitute (4.12) in (4.11) to get the following set of recursive equations for the path probability $P^{(\ell)}_{i,0..t}$ on the graph $G^{(\ell)}$

$$\begin{aligned} P^{(\ell)}_{i,0..t}\left(\sigma^{0..t}|\theta^{1..t}\right) &= \sum_{\sigma_{\partial_i^{\rm in}\setminus\ell}^{0..t-1}}\left(\prod_{j\in\partial_i^{\rm in}\setminus\ell}P^{(i)}_{j,0..t-1}\left(\sigma_j^{0..t-1}|\epsilon_{ij}\zeta_j^{(i),1..t-1}\right)\right) \\ &\quad\times\left(\prod_{s=1}^{t}W\left[\sigma(s)|h_i^{(\ell)}(s)\right]p_0(\sigma(0))\right), \end{aligned} \tag{4.13}$$

with $\ell\in\partial i$. Analogously to (4.4) we have defined

$$h_i^{(\ell)}(s)\equiv\sum_{j\in\partial_i^{\rm in}\setminus\ell}h_{j\to i}(\sigma_j(s-1))+\theta(s). \tag{4.14}$$

To derive (4.13) we have used $P^{(i,\ell)}_{j,0..t}=P^{(i)}_{j,0..t}$. For clarity we have introduced the prefactor $\epsilon_{ij}=c_{ij}c_{ji}$. This factor determines whether or not the edge is symmetric:

$\epsilon_{ij} = 1$ for undirected edges and $\epsilon_{ij} = 0$ for directed edges. The set of $|E|$-equations (4.13) determine the $|E|$-probability distributions $P^{(\ell)}_{i,0..t}\left(\sigma^{0..t}|\theta_i^{1..t}\right)$ at time step t as a function of the $|E|$-probability distributions $P^{(\ell)}_{i,0..t-1}(\sigma^{0..t-1}|\theta_i^{1..t-1})$ at the previous time step $t-1$. In equation (4.13) we only need to take the product over $j \in \partial_i^{\text{in}}$ because the fields $h_i^{(\ell)}(s)$ depend only on $\sigma_{\partial_i^{\text{in}}}$. We call the equations (4.13) the dynamical cavity equations analogously to the static equations (2.88). To find the marginal distributions $P_{i,0..t}$ on the original graph G from the cavity distributions, we need to combine equations (4.11) and (4.12):

$$\begin{aligned} P_{i,0..t}(\sigma^{0..t}|\theta^{1..t}) \quad &= \sum_{\sigma^{0..t-1}_{\partial_i^{\text{in}}}} \left(\prod_{j\in\partial_i^{\text{in}}} P^{(i)}_{j,0..t-1}\left(\sigma_j^{0..t-1}|\epsilon_{ij}\zeta_j^{(i),1..t-1}\right)\right) \\ &\quad \times \left(\prod_{s=1}^{t} W\left[\sigma(s)|h_i(s)\right] p_0(\sigma(0))\right). \end{aligned} \tag{4.15}$$

Equations (4.15) are the dynamical versions of the static equations (2.89).

The main difference between the BP equations (2.88) and the dynamic equations (4.13) is that the latter are iterative equations. Both set of equations have a linear complexity $\mathcal{O}(N)$ in the system size N. However, the iterative nature of the dynamical MPA makes them hard to solve when t becomes large. Indeed, the computational complexity scales $\mathcal{O}(2^t)$ in the time t. The dynamics are only solvable for a very small number of time steps.

The cavity equations simplify a lot when the graph is fully asymmetric. In this case we can set $\epsilon_{ij} = 0$ in equation (4.13). Therefore the equations only have to be solved for $\theta^{0..t} = 0^{0..t}$, where $0^{0..t}$ is the null vector. Moreover, because $\epsilon_{ij} = 0$ the self-coupling disappears in (4.13). We can thus sum on the left and right hand side of (4.13) over $(\sigma_i(0), \sigma_i(1), \cdots, \sigma_i(t-1))$ to get

$$P^{(\ell)}_{i,t}(\sigma) = \sum_{\{\sigma_j\}_{j\in\partial_i^{\text{in}}}} \left(\prod_{j\in\partial_i^{\text{in}}} P^{(i)}_{j,t-1}\left(\sigma_j\right)\right) W\left[\sigma|h\right]. \tag{4.16}$$

We refer to the equations (4.16) as the asymmetric belief propagation algorithm (ABP). The equations (4.15) concern the non-Markovian dynamics of a single effective spin, while the equations (4.16) form a Markovian dynamics of an effective spin. The consequence is that (4.16) is an efficient MPA that determines the marginals of the dynamics of spin models on fully asymmetric graphs with a linear complexity in the time and the system size.

4.3 The ensemble-averaged distribution of paths

We calculate the average of equation (4.13) over all links in the graph, i.e. all $a \in E$. We look at ensembles where the typical graphs have a local tree structure and the degrees on different sites are uncorrelated. We consider graphs drawn from the graph ensembles $\mathcal{G}_{\rm deg}\left(N, p_{\rm d}\left(k^{\rm in}, k^{\rm out}, k^{\rm sym}\right)\right)$ of appendix B.2.5.

The joint degree distribution $p_{\rm d}$ is defined through a histogram as

$$p_{\rm d}(k^{\rm in}, k^{\rm out}, k^{\rm sym}) \equiv \frac{\sum_{i=1}^{N} \delta\left(k^{\rm in}; k_i^{\rm in}\right) \delta\left(k^{\rm out}; k_i^{\rm out}\right) \delta\left(k^{\rm sym}; k_i^{\rm sym}\right)}{N}. \tag{4.17}$$

In equation (4.17) we have used the following notations: the indegree $k_i^{\rm in} = |\partial_i^{\rm in}|$, the outdegree $k_i^{\rm out} = |\partial_i^{\rm out}|$ and the symmetric degree $k_i^{\rm sym} = |\partial_i^{\rm in} \cap \partial_i^{\rm out}|$. From the definitions in subsection 4.2.1 we have

$$\sum_{k^{\rm in}, k^{\rm out}, k^{\rm sym}} k^{\rm in} p(k^{\rm in}, k^{\rm out}, k^{\rm sym}) = \sum_{k^{\rm in}, k^{\rm out}, k^{\rm sym}} k^{\rm out} p(k^{\rm in}, k^{\rm out}, k^{\rm sym}), \tag{4.18}$$

which means that the total number of directed edges in the graph leaving a vertex equals the total number of directed edges entering a vertex.

We define $\overline{P}^{\rm d}$ as the average of the path probabilities $P_{i,t}^{(\ell)}$ over all directed edges (i, ℓ) of E:

$$\begin{aligned} \overline{P}^{\rm d}(\sigma^{0..t}) &\equiv \frac{\sum_{(i,\ell)\in E^{\rm d}} P_{i,0..t}^{(\ell)}(\sigma^{0..t}|0^{1..t})}{|E^{\rm d}|} \\ &= \frac{\sum_{i=1}^{N} \sum_{\ell \in \partial_i^{\rm out} \setminus \partial_i^{\rm sym}} P_{i,0..t}^{(\ell)}(\sigma^{0..t}|0^{1..t})}{|E^{\rm d}|}. \end{aligned} \tag{4.19}$$

The average probability mass function $\overline{P}^{\rm sym}$ is defined as the average of $P_{i,t}^{(\ell)}$ over all links belonging to a undirected edge

$$\begin{aligned} \overline{P}^{\rm sym}\left(\sigma^{0..t}|\theta^{1..t}\right) &\equiv \frac{\sum_{(i,\ell)\in E^{\rm sym}} P_{i,0..t}^{(\ell)}\left(\sigma^{0..t}|\theta^{1..t}\right)}{|E^{\rm sym}|} \\ &= \frac{\sum_{i=1}^{N} \sum_{\ell \in \partial_i^{\rm sym}} P_{i,0..t}^{(\ell)}(\sigma^{0..t}|0^{1..t})}{|E^{\rm sym}|}. \end{aligned} \tag{4.20}$$

When we use the property that the spins in the neighborhood of i are uncorrelated, we can write

$$\overline{\prod_{j\in\partial_i\setminus\ell} P^{(i)}_{j,0..t-1}\left(\sigma_j^{0..t-1}|\zeta_j^{(i),1..t-1}\right)} = \prod_{j\in\partial_i\setminus\ell} \overline{P}\left(\sigma_j^{0..t-1}|\zeta_j^{(i),1..t-1}\right). \tag{4.21}$$

It is useful to focus on a specific example. We consider fields of the type

$$h_{j\to i}(s-1) = J_{ji}\sigma_i(s-1), \tag{4.22}$$

where the interactions strengths J_{ij} are i.i.d.r.v. drawn from an arbitrary distribution $P_I(J)$. When we take the average of the update equations (4.13) according to the definitions (4.19) and (4.20) and use (4.21), we find the recursive equations for the averaged probability mass function of paths. These recursive equations are given by:

$$\begin{aligned}
\overline{P}^{\mathrm{d}}\left(\sigma^{0..t}\right) = &\sum_{k^{\mathrm{out}}\geq 0}^{\infty}\sum_{k^{\mathrm{in}}\geq 0}^{\infty}\sum_{k^{\mathrm{sym}}=0}^{\min(k^{\mathrm{out}},k^{\mathrm{in}})} \frac{p_{\mathrm{d}}(k^{\mathrm{in}},k^{\mathrm{out}},k^{\mathrm{sym}})(k^{\mathrm{out}}-k^{\mathrm{sym}})}{c_{\mathrm{out}}-c_{\mathrm{sym}}} \\
&\times \prod_{\ell=k^{\mathrm{sym}}+1}^{k^{\mathrm{in}}} \int dJ_\ell P_I(J_\ell) \sum_{\sigma_\ell^{0..t-1}} \overline{P}^{\mathrm{d}}\left(\sigma_\ell^{0..t-1}\right) \\
&\times \prod_{\ell=1}^{k^{\mathrm{sym}}} \int dJ_\ell P_I(J_\ell) \sum_{\sigma_\ell^{0..t-1}} \overline{P}^{\mathrm{sym}}\left(\sigma_\ell^{0..t-1}|J_\ell\sigma^{0..t-2}\right) \\
&\times p_0(\sigma(0)) \prod_{s\geq 0}^{t-1} W\left[\sigma(s+1)|\theta(s) + \sum_{0<\ell'\leq k^{\mathrm{in}}} J_{\ell'}\sigma_{\ell'}(s)\right],
\end{aligned} \tag{4.23}$$

and

$$\overline{P}^{\rm sym}\left(\sigma^{0..t}|\theta^{1..t}\right) = \sum_{k^{\rm in}\geq 0}^{\infty}\sum_{k^{\rm sym}=0}^{k^{\rm in}} \frac{p_{\rm d}(k^{\rm in},k^{\rm sym})k^{\rm sym}}{c_{\rm sym}}$$

$$\times \prod_{\ell=k^{\rm sym}}^{k^{\rm in}-1}\int dJ_\ell P_I(J_\ell)\sum_{\sigma_\ell^{0..t-1}}\overline{P}^{\rm d}\left(\sigma_\ell^{0..t-1}\right)$$

$$\times \prod_{\ell=1}^{k^{\rm sym}-1}\int dJ_\ell P_I(J_\ell)\sum_{\sigma_\ell^{0..t-1}}\overline{P}^{\rm sym}\left(\sigma_\ell^{0..t-1}|J_\ell\sigma^{0..t-2}\right)$$

$$\times p_0(\sigma(0))\prod_{s\geq 0}^{t-1} W\left[\sigma(s+1)|\theta(s+1)+\sum_{0<\ell'\leq k^{\rm in}-1} J_{\ell'}\sigma_{\ell'}(s)\right]. \qquad (4.24)$$

We have introduced the average connectivities

$$c_{\rm sym} \equiv \sum_{k^{\rm in},k^{\rm out},k^{\rm sym}} p_{\rm d}(k^{\rm in},k^{\rm out},k^{\rm sym})k^{\rm sym}, \qquad (4.25)$$

$$c_{\rm out} \equiv \sum_{k^{\rm in},k^{\rm out},k^{\rm sym}} p_{\rm d}(k^{\rm in},k^{\rm out},k^{\rm sym})k^{\rm out}. \qquad (4.26)$$

We remark that the factor $k^{\rm out}-k^{\rm sym}$ in (4.23) and the factor $k^{\rm sym}$ in (4.24) follow from the sums over the neighborhoods $\partial_i^{\rm out}\setminus\partial_i^{\rm sym}$ and $\partial_i^{\rm sym}$ in, respectively, (4.19) and (4.20).

The averaged probability mass function $\overline{P}^{\rm real}(\sigma^{0..t})$ over the marginals $P_i(\sigma^{0..t})$, defined through $\overline{P}^{\rm real}(\sigma^{0..t}) \equiv \sum_i P_i(\sigma^{0..t})/N$, can be calculated from (4.15):

$$\begin{aligned}
\overline{P}^{\rm real}\left(\sigma^{0..t}|\theta^{1..t}\right) = & \sum_{k^{\rm in}\geq 0}^{\infty}\sum_{k^{\rm sym}=0}^{k^{\rm in}} p_{\rm d}(k^{\rm in},k^{\rm sym}) \\
& \times \prod_{\ell=k^{\rm sym}+1}^{k^{\rm in}} \int dJ_\ell P_I(J_\ell) \sum_{\sigma_\ell^{0..t-1}} \overline{P}^{\rm d}\left(\sigma_\ell^{0..t-1}\right) \\
& \times \prod_{\ell=1}^{k^{\rm sym}} \int dJ_\ell P_I(J_\ell) \sum_{\sigma_\ell^{0..t-1}} \overline{P}^{\rm sym}\left(\sigma_\ell^{0..t-1}|J_\ell\sigma^{0..t-2}\right) \\
& \times p_0(\sigma(0)) \prod_{s\geq 0}^{t-1} W\left[\sigma(s+1)|\theta(s+1) + \sum_{0<\ell'\leq k^{\rm in}} J_{\ell'}\sigma_{\ell'}(s)\right]. \qquad (4.27)
\end{aligned}$$

The Markovian dynamics of N spins defined in (4.3) is reduced to an effective non-Markovian dynamics of one single spin given by the recursion formulas (4.23), (4.24) and (4.27). Equations analogous to (4.23) and (4.24) have been derived in [110] in the context of LDGM channel coding using the generating functional analysis.

For fully asymmetric graphs, see (4.16), we have that $P_{i,t}^{(\ell)}(\sigma) = P_{i,t}(\sigma)$, but the averages, $\overline{P}_t^{\rm d}(\sigma) \equiv \overline{P_{i,t}^{(\ell)}(\sigma)}$ and $\overline{P}_t^{\rm real}(\sigma) \equiv \overline{P_{i,t}(\sigma)}$, over respectively the links and the sites are different. Indeed:

$$\begin{aligned}
\overline{P}_t^{\rm d}(\sigma) = & \sum_{k^{\rm in}} p_{\rm d}(k^{\rm in}) \frac{c(k^{\rm in})}{c_{\rm out}} \prod_{0<\ell\leq k^{\rm in}} \int dJ_\ell P_I(J_\ell) \\
& \times \sum_{\sigma_\ell} \overline{P}_{t-1}^{\rm d}(\sigma_\ell) W\left[\sigma|\sum_{\ell=1}^{k^{\rm in}} J_\ell\sigma_\ell\right], \qquad (4.28)
\end{aligned}$$

$$\begin{aligned}
\overline{P}_t^{\rm real}(\sigma) = & \sum_{k^{\rm in}} p_{\rm d}(k^{\rm in}) \prod_{0<\ell\leq k^{\rm in}} \int dJ_\ell P_I(J_\ell) \\
& \times \sum_{\sigma_\ell} \overline{P}_{t-1}^{\rm d}(\sigma_\ell) W\left[\sigma|\sum_{\ell=1}^{k^{\rm in}} J_\ell\sigma_\ell\right], \qquad (4.29)
\end{aligned}$$

with $c(k^{\rm in}) = \sum_{k^{\rm out}} p(k^{\rm out}|k^{\rm in})k^{\rm out}$ and $c_{\rm out} = \sum_{k^{\rm out}} p(k^{\rm out})k^{\rm out}$. We have used

$$p(k^{\rm out}|k^{\rm in}) = \begin{cases} \frac{p(k^{\rm out},k^{\rm in})}{p_{\rm d}(k^{\rm in})} & \text{when } p_{\rm d}(k^{\rm in}) \neq 0 \\ 0 & p_{\rm d}(k^{\rm in}) = 0 \end{cases} . \tag{4.30}$$

4.4 Examples of dynamics

In this section we define the type of dynamics that we study by specifying the form of the transition probabilities $W[\sigma|h]$ used in equation (4.3).

4.4.1 Glauber dynamics

Under Glauber dynamics every spin $\sigma_i(t) \in \{-1, 1\}$ evolves under the influence of the field $h_i(t-1)$ with a transition probability $W_{\rm g}(\sigma_i(t)|h_i(t))$ defined through:

$$W_{\rm g}[\sigma|h] \equiv \frac{\exp(\beta\sigma h)}{2\cosh(\beta h)}. \tag{4.31}$$

The parameter β is the inverse of the temperature T. It is possible to implement the dynamics defined by (4.31) and (4.3) with the heat-bath algorithm [15]. When the graph is fully symmetric, detailed balance is satisfied with a 2-spin Hamiltonian [126].

As an example we take a look at the recursive equations (4.23), (4.24) and (4.27) on the Poissonian ensemble $\mathcal{G}_p(N, c, \epsilon)$ defined in B.2.4. For $N \to \infty$, this ensemble becomes equivalent to the ensemble $\mathcal{G}_{\rm deg}\left(N, p_{\rm p}\left(k^{\rm in}, k^{\rm out}, k^{\rm sym}\right)\right)$ with

$$\begin{aligned} p_{\rm p}(k^{\rm in}, k^{\rm out}, k^{\rm sym}) \quad &= \left(\exp\left[-c\epsilon\right] \frac{(c\epsilon)^{k^{\rm sym}}}{k^{\rm sym}!}\right) \left(\exp\left[-(1-\epsilon)c\right] \frac{(c(1-\epsilon))^{k^{\rm in}-k^{\rm sym}}}{(k^{\rm in}-k^{\rm sym})!}\right) \\ &\times \left(\exp\left[-(1-\epsilon)c\right] \frac{(c(1-\epsilon))^{k^{\rm out}-k^{\rm sym}}}{(k^{\rm out}-k^{\rm sym})!}\right). \end{aligned} \tag{4.32}$$

Substitution of (4.32) in (4.24) and (4.27) leads to,

$$\overline{P}^{\rm real}(\sigma^{0..t}|\theta^{1..t}) = \overline{P}^{\rm sym}(\sigma^{0..t}|\theta^{1..t}) = \sum_{k\geq 0}\frac{e^{-c}c^k}{k!}\int\left(\prod_{\ell=1}^{k}dJ_\ell P_I(J_\ell)\right)$$

$$\times\prod_{\ell=1}^{k}\sum_{\sigma_\ell^{0..t-1}}\left[\epsilon\overline{P}^{\rm sym}\left(\sigma_\ell^{0..t-1}|J_\ell\sigma^{0..t-2}\right)+(1-\epsilon)\overline{P}^{\rm sym}(\sigma_\ell^{0..t-1}|0^{0..t-1})\right]$$

$$\times p(\sigma(0))\prod_{s\geq 0}^{t-1}\frac{\exp\left[\beta\sigma(s+1)\left(\theta(s+1)+\sum_{0<\ell'\leq k}J_{\ell'}\sigma_{\ell'}(s)\right)\right]}{2\cosh\left[\beta\left(\theta(s+1)+\sum_{0<\ell'\leq k}J_{\ell'}\sigma_{\ell'}(s)\right)\right]}. \qquad (4.33)$$

For the Poissonian ensemble equation (B.18) (see appendix B) is valid such that $\overline{P}^{d}(\sigma^{0..t}) = \overline{P}^{\rm sym}(\sigma^{0..t}|0^{1..t-1}) = \overline{P}^{\rm real}(\sigma^{0..t})$. The equations (4.23), (4.24) and (4.27) reduce to the ensemble-average over the Poissonian ensemble given by the result equation (4.33) of [74].

4.4.2 Coupled dynamics

We define a dynamics of two sets of N spins under the influence of the same thermal noise. The dynamic variables are $(\sigma_i(t),\tau_i(t)) \in \{-1,1\}\times\{-1,1\}$. The spins $(\sigma_i(t),\tau_i(t))$ only feel the influence of their neighbouring spins (σ_i,τ_i), with $i\in\partial_i^{\rm in}$, through, respectively, the fields $h_i(t)$ and $g_i(t)$. The fields $h_i(t)$ and $g_i(t)$ depend, respectively, only on the sets $\{\sigma_i(t-1)\}_{i=1..N}$ and $\{\tau_i(t-1)\}_{i=1..N}$. The spins $(\sigma_i(t),\tau_i(t))$ evolve according to $W_{\rm c}\left[(\sigma_i(t),\tau_i(t));h_i(t),g_i(t)\right]$:

$$W_{\rm c}\left[(\sigma,\tau)|h,g\right] \equiv$$

$$\delta\left(\sigma;-\tau\right)|r_h-r_g|\left[\Theta\left(r_h-r_g\right)\delta\left(\sigma;1\right)\delta\left(\tau;-1\right)+\Theta\left(r_g-r_h\right)\delta\left(\sigma;-1\right)\delta\left(\tau;1\right)\right]$$

$$+\delta\left(\sigma;\tau\right)\left(1-|r_h-r_g|\right)\Theta(r_h-r_g)\left[\delta(\sigma;1)\frac{r_g}{1+r_g-r_h}+\delta(\sigma;-1)\frac{1-r_h}{1+r_g-r_h}\right]$$

$$+\delta\left(\sigma;\tau\right)\left(1-|r_h-r_g|\right)\Theta(r_g-r_h)\left[\delta(\sigma;1)\frac{r_h}{1+r_h-r_g}+\delta(\sigma;-1)\frac{1-r_g}{1+r_h-r_g}\right]. \qquad (4.34)$$

Θ is the Heaviside step function and the weights r_h and r_g are given by

$$r_h = \frac{\exp\left(\beta h\right)}{2\cosh\left(\beta h\right)},\quad r_g = \frac{\exp\left(\beta g\right)}{2\cosh\left(\beta g\right)}. \qquad (4.35)$$

Equation (4.34) can be simulated using a heat-bath algorithm where at each time step we choose the same random numbers, i.e. the same noise, for both set of spins $\{\sigma\}_{i=1..N}$ and $\{\tau\}_{i=1..N}$. A more compact form of $W_{\rm c}$ is:

$$
\begin{aligned}
&W_{\rm c}\left[(\sigma,\tau)|h,g\right] = \\
&\quad \delta\left(\sigma;-\tau\right)|r_h - r_g|\left[\Theta\left(r_h - r_g\right)\delta\left(\sigma;1\right)\delta\left(\tau;-1\right) + \Theta\left(r_g - r_h\right)\delta\left(\sigma;-1\right)\delta\left(\tau;1\right)\right] \\
&\quad +\delta\left(\sigma;\tau\right)\Theta(r_h - r_g)\left[\delta(\sigma;1)r_g + \delta(\sigma;-1)\left(1 - r_h\right)\right] \\
&\quad +\delta\left(\sigma;\tau\right)\Theta(r_g - r_h)\left[\delta(\sigma;1)r_h + \delta(\sigma;-1)\left(1 - r_g\right)\right]. \qquad (4.36)
\end{aligned}
$$

We use the transition probability $W_{\rm c}$ to determine the phase transitions to a chaotic phase. When the thermal average of the distance between the paths $\{\sigma(t)\}_{i=1..N}$ and $\{\tau_i(t)\}_{i=1..N}$ does not converge to zero for $t \to \infty$, even when the initial distance between $\{\sigma_i(0)\}_{i=1..N}$ and $\{\tau_i(0)\}_{i=1..N}$ is very small, the system is said to be in a chaotic phase. This chaotic behaviour has been studied in [39] for spin glasses and in [40] for neural networks. The coupled dynamics (4.36) does not satisfy detailed balance.

4.5 The path entropy

The fluctuations of the path probabilities $P^{(\ell)}_{i,0..t}$ over all links are given by the distribution of the probabilities of the paths which we will call $\mathcal{P}$. They determine quantities like the average entropy rate $\overline{\mathcal{S}}(t)$. With the recursive equations for the distributions $\mathcal{P}$, we discuss in section 4.10 the stationary solutions of the dynamics.

The stationary entropy rate of a Markov process is defined as [32]

$$
\lim_{t\to\infty}\frac{\overline{\mathcal{S}}(t)}{t} \equiv -\lim_{t\to\infty} t^{-1}\overline{\sum_{\boldsymbol{\sigma}^{0..t}} P_{0..t}(\boldsymbol{\sigma}^{0..t})\log\left(P_{0..t}(\boldsymbol{\sigma}^{0..t})\right)}, \qquad (4.37)
$$

where the bar denotes the average over the quenched variables. With the cavity method we can derive, analogously as done in subsection 2.3.1

$$
\mathcal{S}_G(t) = \sum_{i=1}^{N}\Delta\mathcal{S}_i^{\rm site}(t) - \frac{1}{2}\sum_{a\in E^{\rm sym}}\Delta\mathcal{S}_a^{\rm link}(t) - \sum_{a\in E^{\rm d}}\Delta\mathcal{S}_a^{\rm link}(t). \qquad (4.38)
$$

The quantity $\Delta\mathcal{S}_i^{\rm site}(t)$ is the increment in the entropy $\mathcal{S}$ when the i-th site is added to the graph $G^{(i)}$:

$$\Delta\mathcal{S}_i^{\rm site}(t) = -\sum_{\sigma^{0..t}} \sum_{\sigma_{\partial_i}^{0..t}} P_{\partial_i,0..t}(\sigma_{\partial_i}^{0..t}) \prod_{s=1}^{t} W\left[\sigma(s)|h(s)\right] p_0(\sigma(0)) \\ \times \log\left(P_{\partial i,0..t}(\sigma_{\partial_i}^{0..t}) \prod_{s=1}^{t} W\left[\sigma(s)|h(s)\right] p_0(\sigma(0)) \right), \tag{4.39}$$

with

$$P_{\partial_i,0..t}(\sigma_{\partial_i}^{0..t}) = \prod_{(j,i)\in E^{\rm d}} P_{j,0..t}^{(i)}(\sigma_j^{0..t}) \prod_{(j,i)\in E^{\rm sym}} P_{j,0..t}^{(i)}(\sigma_j^{0..t}|J_{ij}\sigma^{0..t-1}). \tag{4.40}$$

The quantity $\Delta\mathcal{S}_a^{\rm link}(t)$ is minus the entropy difference when one removes the link a from the graph G

$$\Delta\mathcal{S}_{(i,j)}^{\rm link}(t) = -\sum_{\sigma^{0..t},\tau^{0..t}} P_{i,0..t}^{(j)}(\sigma^{0..t}|J_{ji}\tau^{0..t-1}) P_{j,0..t}^{(i)}(\tau^{0..t}|J_{ij}\sigma^{0..t-1}) \\ \times \log\left(P_{i,0..t}^{(j)}(\sigma^{0..t}|J_{ji}\tau^{0..t-1}) P_{j,0..t}^{(i)}(\tau^{0..t}|J_{ij}\sigma^{0..t-1}) \right). \tag{4.41}$$

The summation over the sites in equation (4.38) can be made when we know the distributions of the probabilities of paths $\mathcal{P}$ on the graph.

4.6 The distribution of the probability distributions of paths

We define the following distributions

$$\mathcal{P}^{\rm d}(P) \equiv \frac{\sum_{(i,\ell)\in E^{\rm d}} \prod_{\sigma^{0..t}} \delta\left(P(\sigma^{0..t}) - P_{i,0..t}^{(\ell)}(\sigma^{0..t}) \right)}{|E^{\rm d}|}, \tag{4.42}$$

$$\mathcal{P}^{\rm sym}(P) \equiv \frac{\sum_{(i,\ell)\in E^{\rm sym}} \prod_{\sigma^{0..t},\theta^{1..t}} \delta\left(P(\sigma^{0..t}|\theta^{1..t}) - P_{i,0..t}^{(\ell)}(\sigma^{0..t}|\theta^{1..t}) \right)}{|E^{\rm sym}|},$$

$$\mathcal{P}^{\rm real}(P) \equiv \frac{\sum_{i=1}^{N} \prod_{\sigma^{0..t}} \delta\left(P(\sigma^{0..t}) - P_i(\sigma^{0..t}) \right)}{N}. \tag{4.43}$$

When the variables J_{ij} are i.i.d.r.v. and the transition probabilities are governed by Glauber dynamics, $\mathcal{P}^{\rm d}$ satisfies the recursive equation:

$$\mathcal{P}^{\rm d}\left(P\right) = \sum_{k^{\rm out}\geq 0}^{\infty}\sum_{k^{\rm in}\geq 0}^{\infty}\sum_{k^{\rm sym}=0}^{\min(k^{\rm out},k^{\rm in})}\frac{p_{\rm d}(k^{\rm in},k^{\rm out},k^{\rm sym})(k^{\rm out}-k^{\rm sym})}{c_{\rm out}-c_{\rm sym}}$$
$$\times\prod_{\ell=1}^{k^{\rm sym}}\int dJ_\ell P_I(J_\ell)\int dP_\ell\mathcal{P}^{\rm sym}\left(P_\ell\right)\prod_{\ell=k^{\rm sym}+1}^{k^{\rm in}}\int dJ_\ell P_I(J_\ell)\int dP_\ell\mathcal{P}^{\rm out}\left(P_\ell\right)$$
$$\times\prod_{\sigma^{0..t}}\delta\Bigg(P(\sigma^{0..t})-\mathcal{F}^{\rm d}\left(\sigma^{0..t};\{P_{\ell'},J_{\ell'}\}_{\ell'=1..k^{\rm in}}\right)\Bigg), \tag{4.44}$$

with

$$\mathcal{F}^{\rm d}\left(\sigma^{0..t};\{P_\ell,J_\ell\}_{\ell=1..k^{\rm in}}\right) = \sum_{\sigma_1^{0..t-1},\cdots,\sigma_{k^{\rm in}}^{0..t-1}}\prod_{\ell=1}^{k^{\rm sym}}P_\ell(\sigma_\ell^{0..t-1}|J_\ell\sigma^{0..t-2})$$
$$\times\prod_{\ell=k^{\rm sym}+1}^{k^{\rm in}}P_\ell(\sigma_\ell^{0..t-1})\left(p_0(\sigma(0))\prod_{s\geq 0}^{t-1}\frac{\exp\left[\beta\sigma(s+1)\sum_{0<\ell'\leq k^{\rm in}}J_{\ell'}\sigma_{\ell'}(s)\right]}{2\cosh\left[\beta\sum_{0<\ell'\leq k^{\rm in}}J_{\ell'}\sigma_{\ell'}(s)\right]}\right). \tag{4.45}$$

The distribution along symmetric links is given by $\mathcal{P}^{\rm sym}$

$$\mathcal{P}^{\rm sym}\left(P\right) = \sum_{k^{\rm in}\geq 0}^{\infty}\sum_{k^{\rm sym}=0}^{k^{\rm in}}\frac{p_{\rm d}(k^{\rm in},k^{\rm sym})k^{\rm sym}}{c_{\rm sym}}$$
$$\times\prod_{\ell=1}^{k^{\rm sym}-1}\int dJ_\ell P_I(J_\ell)\int dP_\ell\mathcal{P}^{\rm sym}\left(P_\ell\right)\prod_{\ell=k^{\rm sym}}^{k^{\rm in}-1}\int dJ_\ell P_I(J_\ell)\int dP_\ell\mathcal{P}^{\rm out}\left(P_\ell\right)$$
$$\times\prod_{\sigma^{0..t},\theta^{1..t}}\delta\Bigg(P\left(\sigma^{0..t}|\theta^{1..t}\right)-\mathcal{F}^{\rm sym}\left(\sigma^{0..t}|\theta^{1..t};\{P_{\ell'},J_{\ell'}\}_{\ell'=1..k^{\rm in}-1}\right)\Bigg), \tag{4.46}$$

with

$$\mathcal{F}^{\rm sym}\left(\sigma^{0..t}|\theta^{1..t};\{P_\ell,J_\ell\}_{\ell=1..k^{\rm in}-1}\right)$$
$$=\sum_{\sigma_1^{0..t-1},\cdots,\sigma_{k^{\rm in}-1}^{0..t-1}}\prod_{\ell=1}^{k^{\rm sym}-1}P_\ell(\sigma_\ell^{0..t-1}|J_\ell\sigma_\ell^{0..t-2})\prod_{\ell=k^{\rm sym}}^{k^{\rm in}-1}P_\ell(\sigma_\ell^{0..t-1})$$
$$\times\left(p_0(\sigma(0))\prod_{s\geq 0}^{t-1}\frac{\exp\left[\beta\sigma(s+1)\left(\theta(s+1)+\sum_{0<\ell'\leq k^{\rm in}-1}J_{\ell'}\sigma_{\ell'}(s)\right)\right]}{2\cosh\left[\beta\theta(s+1)+\beta\sum_{0<\ell'\leq k^{\rm in}-1}J_{\ell'}\sigma_{\ell'}(s)\right]}\right). \tag{4.47}$$

The distribution of the single site marginals P_i on the original graph is given by

$$\mathcal{P}(P)=\sum_{k^{\rm in}\geq 0}^{\infty}\sum_{k^{\rm sym}=0}^{k^{\rm in}}p_{\rm d}(k^{\rm in},k^{\rm sym})$$
$$\times\int\prod_{\ell=1}^{k^{\rm sym}}dJ_\ell P_I(J_\ell)\int dP_\ell\mathcal{P}^{\rm sym}(P_\ell)\prod_{\ell=k^{\rm sym}+1}^{k^{\rm in}}\int dJ_\ell P_I(J_\ell)\int dP_\ell\mathcal{P}^{\rm d}(P_\ell)$$
$$\times\prod_{\sigma^{0..t}}\delta\left(P(\sigma^{0..t})-\mathcal{F}^{\rm real}\left(\sigma^{0..t};\{P_{\ell'},J_{\ell'}\}_{\ell'=1..k^{\rm in}}\right)\right), \tag{4.48}$$

with

$$\mathcal{F}^{\rm real}\left(\sigma^{0..t};\{P_\ell,J_\ell\}_{\ell=1..k^{\rm in}}\right)=\sum_{\sigma_1^{0..t-1},\cdots,\sigma_{k^{\rm in}}^{0..t-1}}\prod_{\ell=1}^{k^{\rm sym}}P_\ell(\sigma_\ell^{0..t-1}|J_\ell\sigma_\ell^{0..t-2})$$
$$\times\prod_{\ell=k^{\rm sym}+1}^{k^{\rm in}}P_\ell(\sigma_\ell^{0..t-1})\left(p_0(\sigma(0))\prod_{s\geq 0}^{t-1}\frac{\exp\left[\beta\sigma(s+1)\sum_{0<\ell'\leq k^{\rm in}}J_{\ell'}\sigma_{\ell'}(s)\right]}{2\cosh\left[\beta\sum_{0<\ell'\leq k^{\rm in}}J_{\ell'}\sigma_{\ell'}(s)\right]}\right). \tag{4.49}$$

When we compare the equations (4.44) and (4.46) with the density evolution equations (2.35) for the statics, we see a couple of differences. Since the graph is directed we have now two distributions: one for the probabilities propagating along symmetric edges and one for the probabilities propagating along directed edges.

Since the equations (4.44) and (4.46) describe the dynamics of the model they are recursive equations and their computational complexity scales exponentially in time. Equation (4.48) for the marginals of the path probability is the dynamical equivalent of (2.37) for the marginals of the stationary distribution.

4.7 Comparison with simulations

In this section we compare the value of the magnetisation $m(t)$

$$m(t) = \sum_{\sigma^{0..t}} \sigma(t) \overline{P}^{\text{real}} \left(\sigma^{0..t} | 0^{1..t} \right), \tag{4.50}$$

predicted by equations (4.23), (4.24) and (4.27), with simulation results. It is difficult to develop an efficient Eisfeller-Opper scheme [47] for these equations because the probability distributions of the paths $\overline{P}^{\text{sym}}$ depend on the fields $\theta^{1..t}$. It is necessary to solve (4.24) for all 2^t possible values of $\theta^{1..t}$. We calculate the first time steps through an exact enumeration of the summations over the spin values in equations (4.23), (4.24) and (4.27). The random graphs are drawn according to the algorithm described in subsection B.2.5 of appendix B.

In figures 4.1 and 4.2 we compare the magnetisation in the first time steps, calculated through an exact enumeration, with the results obtained through Monte Carlo simulations. In figure 4.1 we present the results for a bond disordered Ising model on a symmetric Bethe lattice. The interactions J are drawn from the bimodal distribution P_I:

$$P_I(J) = \left(\frac{1+\rho}{2} \right) \delta(J - J_0) + \left(\frac{1-\rho}{2} \right) \delta(J + J_0), \tag{4.51}$$

with ρ the bias in the couplings. In figure 4.2 we show results for the Ising model on a Bethe lattice without bond disorder for a different degree of asymmetry. Both methods are in agreement confirming our claim that the equations (4.23), (4.24) and (4.27) describe exactly the parallel dynamics of spin systems on random graphs.

4.8 The Ising model on a fully asymmetric Bethe lattice

The dynamical cavity equations (4.13) simplify to (4.16) for fully asymmetric graphs. We illustrate the usefulness of these equations for the Ising model on

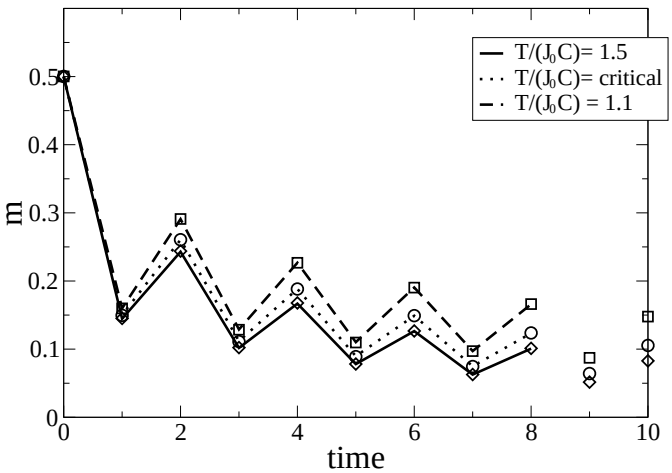


Figure 4.1: The magnetisation m as a function of the discrete time for an Ising model on a symmetric Bethe lattice with connectivity $C = 3$. The interactions are drawn from the bimodal distribution (4.51) with $\rho = 0.25$. The exact enumeration of the recursive equations (4.24) and (4.27) (lines connecting markers) are compared with Monte Carlo simulations (markers).

a graph with a degree distribution $p_{\rm d}(k^{\rm in}, k^{\rm out}, k^{\rm sym}) = \delta(k^{\rm in} - C)p_{\rm d}(k^{\rm out})\delta(k^{\rm sym})$. We call the graphs drawn from this ensemble asymmetric Bethe lattices. The distribution of the outdegrees $p_{\rm d}(k^{\rm out})$ will have no influence on the dynamics, as explicitly seen in equations (4.28) and (4.29). We can therefore chose any arbitrary normalized distribution with $\sum_{k^{\rm out}} k^{\rm out} p_{\rm d}(k^{\rm out}) = C$. Because $c(k^{\rm in}) = c_{\rm out}$ in equation (4.28), $\overline{P}_t^{\rm real}(\sigma) = \overline{P}_t^{\rm d}(\sigma)$ and we only need to solve the recursive equation (4.28). We discuss this model for Glauber dynamics and coupled dynamics as defined in section 4.4.

4.8.1 Glauber dynamics

In this subsection we let the spins evolve according to Glauber dynamics defined in subsection 4.4.1. The transition probability is given by equation (4.31) with a field $h_i(s) = \sum_{j \in \partial_i^{\rm in}} J_{ji}\sigma_j(s-1)$.

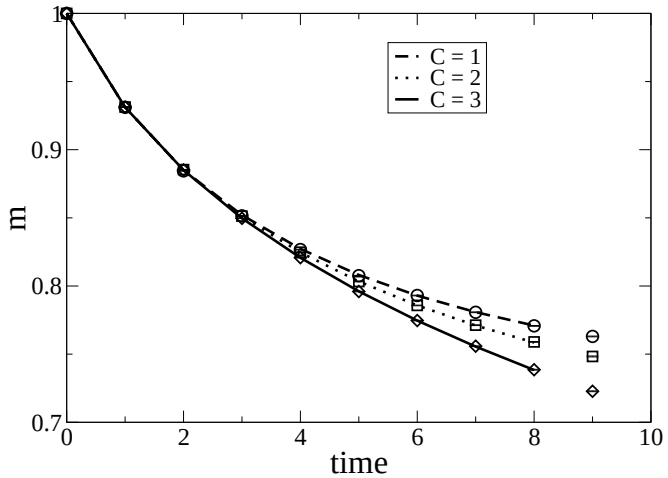


Figure 4.2: The time evolution of the magnetisation m at $T = (1.8J_0)$ for an Ising model on a Bethe lattice without bond disorder for different levels of asymmetry. Results are shown for graphs with fixed indegree $k^{\text{in}} = 3$ and a given fixed outdegree $k^{\text{out}} = C$. The exact enumeration of the recursive equations (4.23), (4.24) and (4.27) (lines connecting markers) are compared with Monte Carlo simulations (markers).

Iterative equations

We derive the equations that give the evolution in time of the following macroscopic observables: the average magnetisation, the correlation function and the distribution of magnetisations.

We define the magnetisation $m(t)$ through the relation $\overline{P}_t(\sigma) = (m(t)\sigma + 1)/2$. From equation (4.28) we get for the evolution of the magnetisation under Glauber dynamics

$$m(t+1) = \prod_{0<\ell\leq C} \left[\sum_{\sigma_\ell} \left(\frac{1+\sigma_\ell m(t)}{2} \right) \right] \left\langle \tanh \left[\beta \left(\theta(t) + \sum_{0<\ell\leq C} J_\ell \sigma_\ell \right) \right] \right\rangle_{J_1,\cdots,J_C} \tag{4.52}$$

To find an expression for the correlation function $C(t,t')$ between spins at time

step t and t', we sum equation (4.23) over all spins, except $\sigma(t)$ and $\sigma(t')$. We get the recursive equation for the two time marginal $P_{t,t'}(\sigma(t), \sigma(t'))$. When we define $C(t, t')$ through

$$P_{t,t'}(\sigma, \tau) = \frac{1}{4}\left[1 + m(t)\sigma + \tau m(t') + C(t, t')\sigma\tau\right], \tag{4.53}$$

we get the recursive equation for the correlation function:

$$C(t+1, t'+1) = \prod_{0<\ell\leq C}\left(\sum_{\sigma_\ell\tau_\ell}\frac{[1 + \sigma_\ell m(t) + \tau_\ell m(t') + \sigma_\ell\tau_\ell C(t, t']}{4}\right)$$
$$\left\langle \tanh\left[\beta\sum_{0<\ell\leq C} J_\ell\sigma_\ell\right]\tanh\left[\beta\sum_{0<\ell\leq C} J_\ell\tau_\ell\right]\right\rangle_{J_1, J_2, \cdots, J_C}. \tag{4.54}$$

For the distribution of the marginals $P_{i,t}^{(\ell)}$ in equation (4.16), or equivalently the distribution $W_t(m)$ of the corresponding magnetisations, we get the recursive equation

$$W_t(m) = \int\prod_{\ell=1}^{C}(dJ_\ell P_I(J_\ell))\int\prod_{\ell=1}^{C}(dm_\ell W_{t-1}(m_\ell))$$
$$\delta\left(m - \prod_{0<\ell\leq C}\sum_{\sigma_\ell}\left(\frac{1 + \sigma_\ell m_\ell}{2}\right)\tanh\left[\beta\sum_{0<\ell'\leq C} J_{\ell'}\sigma_{\ell'}\right]\right). \tag{4.55}$$

The time evolutions determined by the equations (4.52), (4.54) and (4.55) are confirmed by numerical simulations.

A stationary solution

Using the equations (4.52), (4.54) and (4.55) we can find the stationary solutions. We consider the stationary solution $m(t+1) = m(t) = m$. This is a stationary solution without cycles. Substitution of this ansatz in (4.52) shows that m is a solution of

$$m = \prod_{0<\ell\leq C}\left[\sum_{\sigma_\ell}\left(\frac{1 + \sigma_\ell m}{2}\right)\right]\left\langle \tanh\left[\beta\left(\sum_{0<\ell'\leq C} J_{\ell'}\sigma_\ell\right)\right]\right\rangle_{J_1, J_2, \cdots, J_C} \tag{4.56}$$

The model has a phase transition between a F-phase with $m > 0$ at low temperatures and a P-phase with $m = 0$ at high temperatures. Since the phase transition is continuous it is possible to determine the P to F phase transition line with an expansion of the right hand side of (4.56) around $m = 0$. The critical inverse temperature β_F between the P phase and F phase is the solution of

$$1 = \rho 2^{-C} \sum_{r=0}^{C} \binom{k}{r} |2r - k| \tanh\left(\beta^* J \, |2r - k|\right). \tag{4.57}$$

Equation (4.57) holds for the bimodal distributions $P_I(J)$ of the form (4.51). Using the stationary ansatz $q = C(t, t') = C(t - n, t' - n)$ in (4.54) we try find a phase transition between a P phase with $q = 0$ and a SG phase with $q > 0$. We find that $q = 0$ for all temperatures and biases ρ, analogously to [74]. In figure 4.3 we show the P to F transitions (solid lines) for different values of the connectivity C as a function of the temperature T and the bias ρ in the couplings .

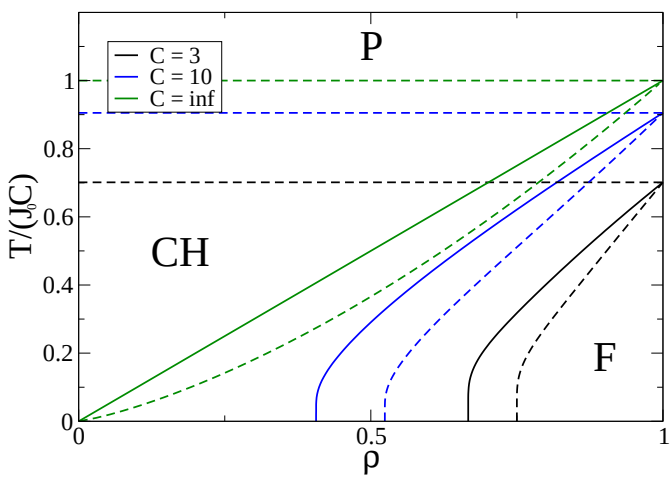


Figure 4.3: The P to F transition lines (solid) for the Ising model on an asymmetric Bethe graph are presented as a function of the rescaled temperature $T/(J_0 C)$ and the bias ρ in the couplings (see equation (4.51)). Phase transition lines for different connectivities C are shown. The dashed lines enclose the regions where the model behaves chaotically (CH).

4.8.2 The chaotic phase

Although the spin model studied in this section has no SG phase, it has a chaotic phase (CH phase). The spins do not freeze into a certain configuration but evolve chaotically in the configuration space. In order to find this phase it is necessary to consider the dynamics of two set of spins $\{\sigma_i\}_{i=1..N}$ and $\{\tau_i\}_{i=1..N}$ that interact on the same graph with the same thermal noise and with a slightly different initial configuration. The transition probability of the spins (σ_i, τ_i) is given by (4.36) with $h_i(s) = \sum_{j\in\partial_i^{\rm in}} J_{ji}\sigma_j(s-1)$ and $g_i(s) = \sum_{j\in\partial_i^{\rm in}} J_{ji}\tau_j(s-1)$.

As done in section 4.8 it is possible to derive the recursive equations for the single time marginals $\overline{P}_t^{\rm d}(\sigma,\tau)$. We define the magnetisation m_t and the thermal average of the Hamming distance d_t between the sets $\{\sigma_i\}_{i=1..N}$ and $\{\tau\}_{i=1..N}$ through

$$\overline{P}_t^{\rm d}(\sigma,\tau) = \frac{1}{4}\left(1 + m_t\sigma + m_t\tau + (1-2d_t)\sigma\tau\right). \tag{4.58}$$

In the case of a bimodal distribution $P_I(J)$ we get for d_t the recursive equation

$$\begin{aligned} d_t = &\sum_{n=0}^{C-1} \binom{k^{\rm in}}{n} (d_{t-1})^{k^{\rm in}-n} (1-d_{t-1})^n \\ &\times \int dxdy \left| \frac{\exp\left(\beta(x+|y|)\right)}{2\cosh\left(\beta(x+|y|)\right)} - \frac{\exp\left(\beta(x-|y|)\right)}{2\cosh\left(\beta(x-|y|)\right)} \right| \\ &\times \sum_{v=0}^{n} \binom{n}{v} \left(\frac{1+\rho m_{t-1}/(1-d_{t-1})}{2}\right)^v \\ &\times \left(\frac{1-\rho m_{t-1}/(1-d_{t-1})}{2}\right)^{n-v} \delta\left(x-2v+n\right) \\ &\times \sum_{w=0}^{k^{\rm in}-n} \binom{k^{\rm in}-n}{w} 2^{n-k^{\rm in}} \delta\left(y-2w+k^{\rm in}-n\right). \end{aligned} \tag{4.59}$$

The time evolution of the magnetisation m_t is given by equation (4.52). In the case of a Gaussian distribution of the couplings we find for d_t the same result as the one derived in [39]. The system is chaotic when the initial configuration has $d_0 \approx 0$ but for large times $t \to \infty$ the Hamming distance satisfies $d_t > 0$.

We substitute the stationary ansatz $d_t = d$ and $m_t = m$ in equation (4.59). When $d > 0$ the system is chaotic. Since the transitions are continuous we can study the bifurcations around the $d = 0$ solution. We find the following equation for the

inverse transition temperature β_{CH} to the CH phase

$$1 = \sum_{k^{\mathrm{out}}\geq 0}^{\infty}\sum_{k^{\mathrm{in}}\geq 0}^{\infty} k^{\mathrm{in}} \sum_{u=0}^{\min(k^{\mathrm{out}},k^{\mathrm{in}})} \frac{p(k^{\mathrm{in}},k^{\mathrm{out}},u)(k^{\mathrm{out}}-k^{\mathrm{sym}})}{c_{\mathrm{out}}-c_{\mathrm{sym}}}$$
$$\times \int dxdy \left| \frac{\exp\left(\beta^*(x+1)\right)}{2\cosh\left(\beta^*(x+1)\right)} - \frac{\exp\left(\beta^*(x-1)\right)}{2\cosh\left(\beta^*(x-1)\right)} \right|$$
$$\times \sum_{v=0}^{k^{\mathrm{in}}-1} \left(\begin{array}{c} k^{\mathrm{in}}-1 \\ v \end{array} \right) \left(\frac{1+\rho m}{2}\right)^{v} \left(\frac{1-\rho m}{2}\right)^{k^{\mathrm{in}}-1-v} \delta\left(x-2v+k^{\mathrm{in}}-1\right). \tag{4.60}$$

For $C \to \infty$ equation (4.60) reduces to $T = 4e^{-2\beta^*\rho m}\left(1+e^{-2\beta^*\rho m}\right)^{-2}$. In figure 4.3 the different phase transitions are shown. For the bias in the couplings ρ large enough and decreasing the temperature starting from a large value, we obtain subsequently the following phases: P phase, CH phase, chaotic F phase and non-chaotic F phase.

4.9 Neural network on a scale-free graph

The interactions between neurons in organisms are mostly asymmetric [175]. Introducing asymmetric couplings in models for neural networks increases the biological realism of the models under study. That is why in [35, 40] the Hopfield model was generalized to include asymmetric couplings.

We add some more realism to the model by defining the neural network on a graph with a given degree distribution. Many real-world networks have a degree distribution of the form $p(k) = ak^{-\gamma}$, with a a normalisation constant [5, 80, 130]. These are called scale-free graphs. One example is the network of brain activity which has scale-free features [46]. In [25] neural networks on scale-free graphs with only symmetric couplings have been studied.

We consider a neural network on a fully asymmetric graph with the following distribution of indegrees and outdegrees [153]:

$$p(k^{\mathrm{in}},k^{\mathrm{out}}) = Aak_{\mathrm{out}}^{-\lambda}\delta\left(k^{\mathrm{in}},k^{\mathrm{out}}\right) + (1-A)a^2k_{\mathrm{in}}^{-\lambda}k_{\mathrm{out}}^{-\lambda}. \tag{4.61}$$

The correlation factor A in (4.61) denotes the fraction of sites where the number of connections entering and leaving the site are equal. This correlation between the degrees, which is found in many real-world networks [153], will turn out to

have much influence on the performance of a scale-free neural network. For fixed λ we will change the average number of interactions by increasing the lower bound b: $p(k^{\text{in}}, k^{\text{out}}) = 0$ for $k^{\text{in}} < b$ and $k^{\text{out}} < b$.

We take the strengths of the interactions J_{ij} according to the Hebb rule:

$$J_{ij} = \frac{1}{p} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu, \tag{4.62}$$

with the $\xi_i^\mu \in \{-1, 1\}$ uncorrelated patterns drawn from the probability distribution

$$q(\xi_i^\mu) = \frac{1}{2} \left(\delta(\xi_i^\mu; 1) + \delta(\xi_i^\mu; -1) \right). \tag{4.63}$$

The network has p patterns $\boldsymbol{\xi}_i = \left(\xi_i^1, \xi_i^2, \cdots, \xi_i^p\right)$ on each site. The J_{ij} are not i.i.d.r.v. variables, therefore, we can not use the equations (4.23) and (4.24).

4.9.1 Glauber dynamics

We first derive the recursive equations for the marginal distributions when the variables evolve according to a Glauber dynamics. From these equations we determine the phase transition from a P phase to a retrieval phase (R phase). In the R phase the network can recover a stored pattern while in the P phase the noise is too large to retrieve a stored pattern from a distorted signal.

To calculate the mean of the cavity equations (4.13) over the quenched variables, it is necessary to define the sublattices $I_{\boldsymbol{\xi}}$: $I_{\boldsymbol{\xi}} \equiv \{i \in V | \boldsymbol{\xi}_i = \boldsymbol{\xi}\}$. The averaged path probabilities $\overline{P}^{\text{d}}_{\boldsymbol{\xi}}(\boldsymbol{\sigma}^{(t)})$ and $\overline{P}^{\text{real}}_{\boldsymbol{\xi}}(\boldsymbol{\sigma}^{(t)})$, on the sublattices $I_{\boldsymbol{\xi}}$ are defined as

$$\overline{P}^{\text{d}}_{\boldsymbol{\xi}}\left(\sigma^{0..t}\right) \equiv \frac{\sum_{i \in I_{\boldsymbol{\xi}}} \sum_{(i,\ell) \in E^{\text{d}}} P_i^{(\ell)}\left(\sigma^{0..t} | 0^{1..t}\right)}{\sum_{i \in I_{\boldsymbol{\xi}}} \sum_{(i,\ell) \in E^{\text{d}}}}, \tag{4.64}$$

$$\overline{P}^{\text{real}}_{\boldsymbol{\xi}}(\sigma^{0..t}) \equiv \frac{\sum_{i \in I_{\boldsymbol{\xi}}} P_i(\sigma^{0..t} | 0^{1..t})}{\sum_{i \in I_{\boldsymbol{\xi}}}}. \tag{4.65}$$

When the graph is drawn from an ensemble defined by a degree distribution of the form $p(k^{\text{in}}, k^{\text{out}}, k^{\text{sym}}) = p(k^{\text{in}}, k^{\text{out}})\delta(k^{\text{sym}})$, such that there are no symmetric

couplings, we get the following recursive equation for $\overline{P}^{\rm d}_{\boldsymbol{\xi}}$

$$\overline{P}^{\rm d}_{\boldsymbol{\xi}}\left(\sigma^{0..t}\right) = \sum_{k^{\rm in}=b}^{\infty} p_{\rm d}(k^{\rm in}) \frac{c_{\rm out}\left(k^{\rm in}\right)}{c_{\rm out}} \prod_{0<\ell\leq k^{\rm in}} \sum_{\sigma_\ell^{0..t-1}} \sum_{\boldsymbol{\xi}_\ell} \frac{\overline{P}^{\rm d}_{\boldsymbol{\xi}_\ell}(\sigma_\ell^{0..t-1})}{2^p}$$

$$\times p_0(\sigma(0)) \prod_{s\geq 0} \frac{\exp\left[\beta\sigma(s+1)\sum_{0<\ell'\leq k^{\rm in}} \frac{\boldsymbol{\xi}\cdot\boldsymbol{\xi}_{\ell'}}{p}\sigma_{\ell'}(s)\right]}{2\cosh\left[\beta\sum_{0<\ell'\leq k^{\rm in}} \frac{\boldsymbol{\xi}\cdot\boldsymbol{\xi}_{\ell'}}{p}\sigma_{\ell'}(s)\right]}, \qquad (4.66)$$

with $c_{\rm out}(k^{\rm in})$ the average number of directed edges leaving a site, given the indegree $k^{\rm in}$: $c_{\rm out}(k^{\rm in}) = \sum_{k^{\rm out}} p(k^{\rm out}|k^{\rm in})k^{\rm out}$. For the averaged path probability on the original graph we get analogously

$$\overline{P}^{\rm real}_{\boldsymbol{\xi}}\left(\sigma^{0..t}\right) = \sum_{k^{\rm in}=b}^{\infty} p_{\rm d}(k^{\rm in}) \prod_{0<\ell\leq k^{\rm in}} \sum_{\sigma_\ell^{0..t-1}} \sum_{\boldsymbol{\xi}_\ell} \frac{\overline{P}^{\rm d}_{\boldsymbol{\xi}_\ell}(\sigma_\ell^{0..t-1})}{2^p}$$

$$\times p_0(\sigma(0)) \prod_{s\geq 0} \frac{\exp\left[\beta\sigma(s+1)\sum_{0<\ell'\leq k^{\rm in}} \frac{\boldsymbol{\xi}\cdot\boldsymbol{\xi}_{\ell'}}{p}\sigma_{\ell'}(s)\right]}{2\cosh\left[\beta\sum_{0<\ell'\leq k^{\rm in}} \frac{\boldsymbol{\xi}\cdot\boldsymbol{\xi}_{\ell'}}{p}\sigma_{\ell'}(s)\right]}. \qquad (4.67)$$

We use the notation $P^{\rm a}_{\boldsymbol{\xi}}(\sigma(t)) = \frac{1}{2}\left(1 + m^{\rm a}_{\boldsymbol{\xi}}(t)\sigma(t)\right)$ with superscript a = d or a = real. The magnetisations $m^{\rm d}_{\boldsymbol{\xi}}$ and $m^{\rm real}_{\boldsymbol{\xi}}$ evolve in time according to

$$m^{\rm a}_{\boldsymbol{\xi}}(t) = \sum_{k^{\rm in}=b}^{\infty} p_{\rm d}(k^{\rm in}) \frac{c_{\rm a}\left(k^{\rm in}\right)}{c_{\rm out}} 2^{-pk^{\rm in}} \prod_{0<\ell\leq k^{\rm in}} \sum_{\sigma_\ell} \sum_{\boldsymbol{\xi}_\ell} \left(\frac{1+\sigma_\ell m^{\rm a}_{\boldsymbol{\xi}_\ell}(t-1)}{2}\right)$$

$$\times \tanh\left(\beta \frac{\sum_{0<\ell'\leq k^{\rm in}} \boldsymbol{\xi}\cdot\boldsymbol{\xi}_{\ell'}\sigma_{\ell'}}{p}\right), \qquad (4.68)$$

with for the average connectivities

$$c_{\rm d}(k^{\rm in}) = c_{\rm out}(k^{\rm in}) = (1-A)c_{\rm out} + Ak^{\rm in}, \qquad (4.69)$$

$$c_{\rm real}(k^{\rm in}) = c_{\rm out}. \qquad (4.70)$$

We simplify the equations (4.68) with the condensed ansatz $m^{\rm a}_{\boldsymbol{\xi}}(t) = \xi^1 m^{\rm a}(t)$. The ansatz assumes that the spins (= neurons) only have a finite overlap with the first

pattern. The overlap $m^{\mathrm{a}}(t)$ evolves according to

$$m^{\mathrm{a}}(t) = \sum_{k^{\mathrm{in}}=b}^{\infty} p_{\mathrm{d}}(k^{\mathrm{in}}) \frac{c_{\mathrm{a}}(k^{\mathrm{in}})}{c_{\mathrm{out}}} M(k^{\mathrm{in}}), \tag{4.71}$$

with

$$M(k^{\mathrm{in}}) = 2^{-(p-1)k^{\mathrm{in}}} \sum_{r=0}^{k^{\mathrm{in}}(p-1)} \sum_{s=0}^{k^{\mathrm{in}}} \left(\begin{array}{c} k^{\mathrm{in}}(p-1) \\ r \end{array} \right) \left(\begin{array}{c} k^{\mathrm{in}} \\ s \end{array} \right)$$
$$\times \left[\frac{1+m^{\mathrm{d}}(t-1)}{2} \right]^{s} \left[\frac{1-m^{\mathrm{d}}(t-1)}{2} \right]^{k^{\mathrm{in}}-s} \tanh\left(\frac{\beta(2s+2r-k^{\mathrm{in}}p)}{p} \right). \tag{4.72}$$

When calculating numerically the sum in the degrees k^{in} in equation (4.71) we have to introduce a cutoff K (see also the discussion at the end of subsection 2.2.3). We bound m^{a} by two values, $m^{\mathrm{a}}_{\mathrm{l}} < m^{\mathrm{a}}$ and $m^{\mathrm{a}}_{\mathrm{u}} > m^{\mathrm{a}}$, with $m^{\mathrm{a}}_{\mathrm{l}}$ and $m^{\mathrm{a}}_{\mathrm{u}}$ defined through

$$m^{\mathrm{a}}_{\mathrm{l}}(t) \equiv \sum_{k^{\mathrm{in}}=b}^{K} p_{\mathrm{d}}(k^{\mathrm{in}}) \frac{c_{\mathrm{a}}(k^{\mathrm{in}})}{c_{\mathrm{out}}} M(k^{\mathrm{in}}), \tag{4.73}$$

$$m^{\mathrm{a}}_{\mathrm{u}}(t) \equiv \sum_{k^{\mathrm{in}}=b}^{K} p_{\mathrm{d}}(k^{\mathrm{in}}) \frac{c_{\mathrm{a}}(k^{\mathrm{in}})}{c_{\mathrm{out}}} M(k^{\mathrm{in}})$$
$$+\mathrm{sign}\left(m^{\mathrm{d}}_{\mathrm{u}}(t-1)\right) \sum_{k^{\mathrm{in}}=K+1}^{\infty} p_{\mathrm{d}}(k^{\mathrm{in}}) \frac{c_{\mathrm{a}}(k^{\mathrm{in}})}{c_{\mathrm{out}}}. \tag{4.74}$$

In equation (4.74) we have used that $M(\infty) = \mathrm{sign}\left(m^{\mathrm{d}}(t-1)\right)$. Since we have a power-law decay of the degree distribution (and not an exponential decay) it is important to take the cutoff K into consideration when we want to know the asymptotic behaviour of the neural network for $K \to \infty$. When the degree distribution is power law the macroscopic observables will converge much slower to its asymptotic value $K = \infty$ in contrast to the case of a Poissonian distribution. For finite K the time evolution of equation (4.73) is confirmed by Monte Carlo simulations.

4.9.2 The retrieval state

When we consider a stationary state without cycles of the form, $m(t) = m(t-1) = m$, we get the following equation for the critical inverse temperature β^{R} of the P to R transition

$$1 = \sum_{k^{\mathrm{in}}} p_{\mathrm{d}}(k^{\mathrm{in}}) \frac{c\left(k^{\mathrm{in}}\right)}{c_{\mathrm{out}}} \mathcal{A}(\beta^{\mathrm{R}}, k^{\mathrm{in}}), \tag{4.75}$$

with

$$\mathcal{A}(\beta^{\mathrm{R}}, k^{\mathrm{in}}) = 2^{-pk^{\mathrm{in}}} \sum_{r=0}^{k^{\mathrm{in}}(p-1)} \sum_{s=0}^{k^{\mathrm{in}}} \left(\begin{array}{c} k^{\mathrm{in}}(p-1) \\ r \end{array} \right) \left(\begin{array}{c} k^{\mathrm{in}} \\ s \end{array} \right)$$
$$\times (2s - k^{\mathrm{in}}) \tanh\left(\frac{\beta^{\mathrm{R}}(2s + 2r - k^{\mathrm{in}}p)}{p} \right). \tag{4.76}$$

We have performed a bifurcation analysis on the equations (4.73) and (4.74) to derive equations (4.75) and (4.76). Below the critical temperature $T_R = 1/\beta^R$, $m^{\mathrm{real}} > 0$ such that the neural network can retrieve a stored pattern from a distorted initial configuration. To solve numerically equation (4.75) we take a distribution $p_{\mathrm{d}}(k^{\mathrm{in}}) = a k_{\mathrm{in}}^{-\gamma}$ with $k^{\mathrm{in}} \in [b, \cdots, K]$ and zero for other values of k^{in}. We introduce the lower critical value $\beta_{\mathrm{l}}^{\mathrm{R}}$, and the upper value $\beta_{\mathrm{u}}^{\mathrm{R}}$ through

$$\begin{aligned} 1 &= \sum_{k^{\mathrm{in}}=b}^{K} p_{\mathrm{d}}(k^{\mathrm{in}}) \frac{c\left(k^{\mathrm{in}}\right)}{c_{\mathrm{out}}} \mathcal{A}(\beta_{\mathrm{l}}^{\mathrm{R}}, k^{\mathrm{in}}), \\ 1 &= \sum_{k^{\mathrm{in}}=b}^{K} p_{\mathrm{d}}(k^{\mathrm{in}}) \frac{c\left(k^{\mathrm{in}}\right)}{c_{\mathrm{out}}} \mathcal{A}(\beta_{\mathrm{u}}^{\mathrm{R}}, k^{\mathrm{in}}) \\ &\quad + \sqrt{\frac{2}{\pi p}} \left(\sum_{k^{\mathrm{in}}=K+1}^{\infty} p_{\mathrm{d}}(k^{\mathrm{in}}) \sqrt{k^{\mathrm{in}}} \frac{c\left(k^{\mathrm{in}}\right)}{c_{\mathrm{out}}} \right). \end{aligned} \tag{4.77}$$

To derive equation (4.77) we have used that $\mathcal{A}(\beta, C)$ has the following asymptotic behaviour for $C \to \infty$

$$\mathcal{A}(\beta, C) \to \sqrt{\frac{2}{\pi p}} \sqrt{C}. \tag{4.78}$$

The asymptotic value of the Riemann sum in the second term of (4.77) can be calculated using a series that converges exponentially in K [19]. In figure 4.4 we

compare how $\beta_{\mathrm{l}}^{\mathrm{R}}$ and $\beta_{\mathrm{u}}^{\mathrm{R}}$ converge to their asymptotic value for $K \to \infty$. The upper bound $\beta_{\mathrm{u}}^{\mathrm{R}}$ clearly converges fast to the asymptotic value. The lower bound $\beta_{\mathrm{l}}^{\mathrm{R}}$ on the other hand converges very slow to its asymptotic value. When we would estimate the critical value β^{R} with $\beta_{\mathrm{l}}^{\mathrm{R}}$, which we usually do for Poissonian graphs, we would find a bad value for the critical temperature β^{R}. Therefore we take T_{u}^{R} as an approximation to the critical temperature in figures 4.5 and 4.6. In these figures we have plotted T^{R} as a function of, respectively, the exponent λ and the correlation factor A. The retrieval phase increases with λ, which is expected since the mean connectivity of the graph also increases with λ. Increasing the correlation A between the indegrees and the outdegrees on scale-free graphs ameliorates the performance of the network. The neural network becomes much more tolerant to noise and can retrieve considerably more patterns when A increases.

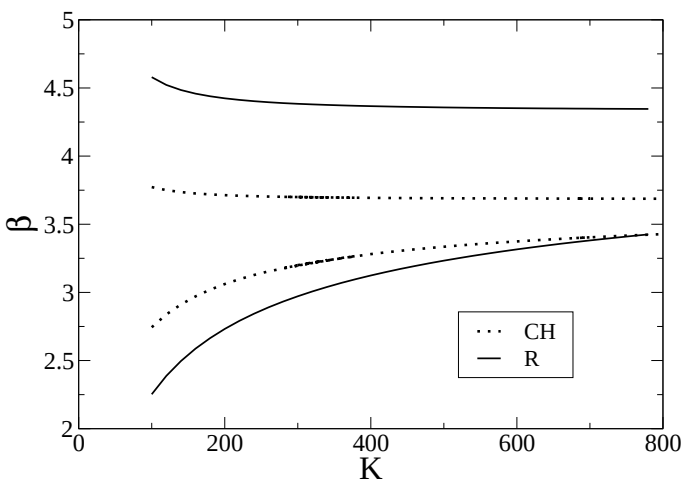


Figure 4.4: Neural network on a scale-free graph: The bounds $\beta_{\mathrm{u}}^{\mathrm{R}}, \beta_{\mathrm{u}}^{\mathrm{ch}}$ (upper lines) and $\beta_{\mathrm{l}}^{\mathrm{R}}, \beta_{\mathrm{l}}^{\mathrm{ch}}$ (lower lines) on the inverse critical temperatures $\beta^{\mathrm{R}}, \beta^{\mathrm{ch}}$ as a function of the cutoff K. The bounds on β^{R} are calculated for the model parameters $\lambda = 2$, $p = 3$, $b = 4$. The bounds on β^{ch} are calculated for $\lambda = 2.2$, $p = 3$, $b = 4$. We see that the upper bounds saturate much faster than the lower bounds.

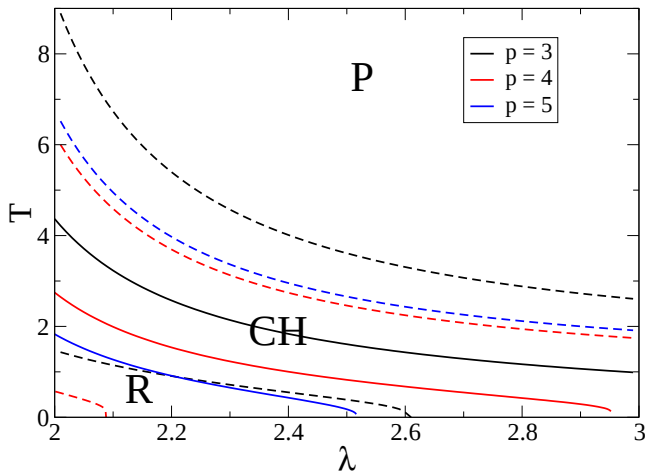


Figure 4.5: Neural network on a scale-free graph: The critical temperatures T^{R} (solid lines) T^{ch} (upper dashed lines) and T^{m} (lower dashed lines) as a function of the exponent λ (see equation (4.61)) for a different number of patterns p. The minimal indegree is $b = 4$ and the correlation factor $A = 0$. The R phase and CH phase become smaller when p increases and the non chaotic part of the R phase dissapears when γ increases. The CH phase is larger for odd values of p than for even values of p.

4.9.3 The chaotic phase

In this subsection we determine the CH phase of the neural network. In [40] this was done on a Poissonian graph using annealed methods where the interactions are drawn at every time step independently.

We consider two systems on the same graph undergoing the same thermal noise through the coupled dynamics described in subsection 4.4.2. We find the following

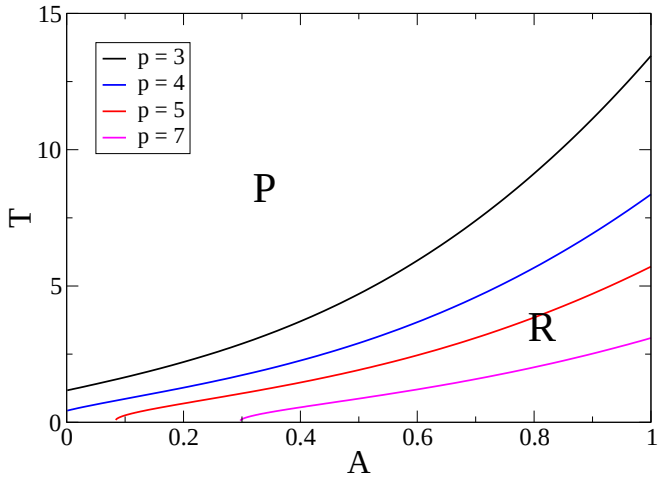


Figure 4.6: The critical temperatures of the retrieval state T^{R} of the neural network as a function of the correlation factor A of the scale-free graph for a different number of patterns p. The ensemble of scale-free graphs has the parameters $\lambda = 2.8$, $b = 4$. The critical temperatures are estimated with $T_{\mathrm{u}}^{\mathrm{R}}$ for $K = 2000$. The R phase increases considerably with the correlation factor A.

Markovian process for the single time marginals $\overline{P}^{\mathrm{d}}_{t,\boldsymbol{\xi}}$:

$$\overline{P}^{\mathrm{d}}_{t,\boldsymbol{\xi}}(\sigma,\tau) = \sum_{k^{\mathrm{in}}} p_{\mathrm{d}}(k^{\mathrm{in}}) \frac{c(k^{\mathrm{in}})}{c_{\mathrm{out}}} \prod_{0<\ell\le k^{\mathrm{in}}} \sum_{\sigma_\ell,\tau_\ell} \sum_{\boldsymbol{\xi}_\ell} \frac{\overline{P}^{\mathrm{d}}_{t-1,\boldsymbol{\xi}_\ell}(\sigma_\ell,\tau_\ell)}{2^p}$$

$$W_{\mathrm{c}}\left[\sigma,\tau\middle|\sum_{0<\ell'\le k^{\mathrm{in}}} \frac{\boldsymbol{\xi}\cdot\boldsymbol{\xi}_{\ell'}}{p}\sigma_{\ell'}, \sum_{0<\ell'\le k^{\mathrm{in}}} \frac{\boldsymbol{\xi}\cdot\boldsymbol{\xi}_{\ell'}}{p}\tau_{\ell'}\right], \tag{4.79}$$

where W_{c} is the transition probability defined in (4.36). We parameterize the single time marginals $\overline{P}^{\mathrm{d}}_{t,\boldsymbol{\xi}}$ with the magnetisations $m_{\boldsymbol{\xi},1}(t)$, $m_{\boldsymbol{\xi},2}(t)$ and the Hamming distance $d_{\boldsymbol{\xi}}(t)$ at time step t:

$$\overline{P}^{\mathrm{d}}_{t,\boldsymbol{\xi}}(\sigma,\tau) = \frac{1}{4}\left[1 + \sigma m_{\boldsymbol{\xi},1}(t) + \tau m_{\boldsymbol{\xi},2}(t) + \sigma\tau\left(1 - 2d_{\boldsymbol{\xi}}(t)\right)\right]. \tag{4.80}$$

The Hamming distance evolves according to

$$d_{\boldsymbol{\xi}}(t) = \sum_{k^{\rm in}=b}^{\infty} p_{\rm d}(k^{\rm in}) \times \prod_{0<\ell\leq k^{\rm in}} \sum_{\sigma_\ell,\tau_\ell} \sum_{\boldsymbol{\xi}_\ell} \left[\frac{1+\sigma_\ell m_{\boldsymbol{\xi}_\ell,1}(t-1)+\tau_\ell m_{\boldsymbol{\xi}_\ell,2}(t-1)+\left(1-2d_{\boldsymbol{\xi}_\ell}(t-1)\right)\sigma_\ell\tau_\ell}{2^{p+2}} \right] \times \left| \frac{\exp\left[\beta\sum_{\ell'}\left(\frac{\boldsymbol{\xi}\cdot\boldsymbol{\xi}_{\ell'}}{p}\right)\sigma_{\ell'}\right]}{2\cosh\left[\beta\sum_{\ell'}\left(\frac{\boldsymbol{\xi}\cdot\boldsymbol{\xi}_{\ell'}}{p}\right)\sigma_{\ell'}\right]} - \frac{\exp\left[\beta\sum_{\ell'}\left(\frac{\boldsymbol{\xi}\cdot\boldsymbol{\xi}_{\ell'}}{p}\right)\tau_{\ell'}\right]}{2\cosh\left]\beta\sum_{\ell'}\left(\frac{\boldsymbol{\xi}\cdot\boldsymbol{\xi}_{\ell'}}{p}\tau_{\ell'}\right)\right]} \right| . \tag{4.81}$$

The magnetisations $m_{\boldsymbol{\xi},1}(t)$ and $m_{\boldsymbol{\xi},2}(t)$ evolve according to equation (4.68).

We use the condensed initial conditions, $m_{\boldsymbol{\xi},1}(0) = \xi^1 m(0)$, $m_{\boldsymbol{\xi},2}(0) = \xi^1 m(0)$ and $d_{\boldsymbol{\xi}}(0) = d(0)$. These initial conditions imply that the initial configurations have a finite overlap with only the first pattern. From the time evolution (4.79) we find that for the condensed initial conditions $m_{\boldsymbol{\xi},1}(t) = \xi^1 m(t)$, $m_{\boldsymbol{\xi},2}(t) = \xi^1 m(t)$ and $d_{\boldsymbol{\xi}}(t) = d(t)$. The evolution of the overlap $m(t)$ is given by equation (4.71). For

$d(t)$ we get

$$
\begin{aligned}
d(t) = & \sum_{k^{\rm in}=b}^{\infty} p_{\rm d}(k^{\rm in}) \sum_{n=0}^{k^{\rm in}} \left(\begin{array}{c} k^{\rm in} \\ n \end{array} \right) (d(t-1))^{k^{\rm in}-n} (1-d(t-1))^{n} \\
& \times \sum_{R_1=0}^{(k^{\rm in}-n)(p-1)} \sum_{R_2=0}^{n(p-1)} f\left(R_1,(p-1)(k^{\rm in}-n)\right) f\left(R_2,(p-1)n\right) \\
& \times \int dxdy \ \exp\left[\frac{\beta}{p}\left(-n(p-1)+2R_2+x\right)\right] \\
& \times \left| \frac{\exp\left[\frac{\beta}{p}\left(|y|+2R_1-(k^{\rm in}-n)(p-1)\right)\right]}{2\cosh\left[\frac{\beta}{p}\left(x+|y|-k^{\rm in}(p-1)+2R_1+2R_2\right)\right]} \right. \\
& \left. \times -\frac{\exp\left[\frac{\beta}{p}\left(-|y|-2R_1+(k^{\rm in}-n)(p-1)\right)\right]}{2\cosh\left[\frac{\beta}{p}\left(x-|y|+(k^{\rm in}-2n)(p-1)-2R_1+2R_2\right)\right]} \right| \\
& \times \sum_{v=0}^{n} \left(\begin{array}{c} n \\ v \end{array} \right) \left(\frac{1+m(t-1)/(1-d(t-1))}{2} \right)^{v} \\
& \times \left(\frac{1-m(t-1)/(1-d(t-1))}{2} \right)^{n-v} \delta\left(x-2v+n\right) \\
& \times \sum_{w=0}^{k^{\rm in}-n} \left(\begin{array}{c} k^{\rm in}-n \\ w \end{array} \right) 2^{-n+k^{\rm in}} \delta\left(y-2w+k^{\rm in}-n\right). \qquad (4.82)
\end{aligned}
$$

In equation (4.82) we have summed subsequently over the following variables: the indegrees $k^{\rm in}$, the number of neighbouring spins n with $\sigma \neq \tau$, the number of neighbouring spins v with $\sigma = \tau$ and $\sigma = 1$ and the number of neighbouring spins w with $\sigma \neq \tau$ and $\sigma = 1$. The summation variables R_1 and R_2 are the number of non-condensed patterns ξ_ℓ^μ on the neighbouring spins that are equal to the corresponding pattern ξ^μ on the original site with, respectively, $\sigma \neq \tau$ and $\sigma = \tau$. The complex function $f(R;x)$ used in equation (4.82) equals

$$
f(R;x) = \frac{1}{2\pi} \int_0^{2\pi} d\omega \exp\left[i\omega R\right] \left|\cos\left(\frac{\omega}{2}\right)\right|^{x} \exp\left[-ix\,\mathrm{atan}\left(\frac{\sin\omega}{1+\cos\omega}\right)\right] \qquad (4.83)
$$

It is possible to find an equation for the inverse transition temperature $\beta^{\rm ch}$ to a

CH phase with the stationary value $d > 0$ of $d(t)$ by expanding the left side of (4.82) around $d(t-1) = 0$:

$$1 = \sum_{k^{\rm in}} p_{\rm d}(k^{\rm in}) k^{\rm in} \frac{c(k^{\rm in})}{c_{\rm out}} \mathcal{B}(\beta, k^{\rm in}, m), \tag{4.84}$$

with

$$\begin{aligned}
\mathcal{B}(\beta, k^{\rm in}, m) = \int dx \sum_{v=0}^{k^{\rm in}-1} & \delta\left(x - 2v + k^{\rm in} - 1\right) \\
& \times \binom{k^{\rm in}-1}{v} \left(\frac{1+m}{2}\right)^{v} \left(\frac{1-m}{2}\right)^{k^{\rm in}-1-v} \\
& \times \sum_{R_1=0}^{p-1} \sum_{R_2=0}^{(k^{\rm in}-1)(p-1)} f\left(R_1, p-1\right) f\left(R_2, (p-1)(k^{\rm in}-1)\right) \\
& \times \exp\left[\frac{\beta}{p}\left(-(k^{\rm in}-1)(p-1) + 2R_2 + x\right)\right] \\
& \times \left| \frac{\exp\left[\frac{\beta}{p}\left(1 + 2R_1 - (p-1)\right)\right]}{2\cosh\left[\frac{\beta}{p}\left((x+1) - k^{\rm in}(p-1) + 2R_1 + 2R_2\right)\right]} \right. \\
& \left. - \frac{\exp\left[\frac{\beta}{p}\left(-1 - 2R_1 + (p-1)\right)\right]}{2\cosh\left[\frac{\beta}{p}\left((x-1) + (-k^{\rm in}+2)(p-1) - 2R_1 + 2R_2\right)\right]} \right|.
\end{aligned} \tag{4.85}$$

The asymptotic behaviour of $\mathcal{B}$ for $C \to \infty$ is given by:

$$\mathcal{B}(\beta, C, m) \sim \exp\left(C\Psi(m)\right), \tag{4.86}$$

with

$$\Psi(m) = \log\left[\frac{(1-m) + (1+m)\exp\left(\frac{2x}{p}\right)}{2}\right] + (p-1)\log\left[\frac{1+\exp\left(\frac{2x}{p}\right)}{2}\right] - x,$$

and

$$x = \frac{p}{2}\log\left[\frac{-\frac{2m}{p} + m + \sqrt{\left(-m+\frac{2m}{p}\right)^2 - (m-1)(1+m)}}{1+m}\right]. \tag{4.87}$$

When $|m| > 0$ we have that $\Psi < 0$, hence, the series converges exponentially for large C. When $m = 0$ we have that $\Psi = 0$. The asymptotic behaviour of $\mathcal{B}$ is then:

$$\mathcal{B}(\beta, C, m) \sim \frac{1}{\sqrt{C}}\left(\sqrt{\frac{2}{p\pi}}\right) 2^{-p+1} \sum_{r=0}^{p-1} \left(\begin{array}{c} p-1 \\ r \end{array} \right) |1 + 2r - (p-1)| = \frac{\zeta}{\sqrt{C}}. \tag{4.88}$$

We define the upper bounds $\beta_{\mathrm{u}}^{\mathrm{ch}}$ and $\beta_{\mathrm{l}}^{\mathrm{ch}}$ on the inverse critical temperature β^{ch} to the CH phase for $m = 0$ as:

$$1 = \sum_{k^{\mathrm{in}}=b}^{K} p_{\mathrm{d}}(k^{\mathrm{in}}) k^{\mathrm{in}} \frac{c(k^{\mathrm{in}})}{c_{\mathrm{out}}} \mathcal{B}(\beta_{\mathrm{l}}^{\mathrm{ch}}, k^{\mathrm{in}}, 0), \tag{4.89}$$

$$\begin{aligned} 1 = & \sum_{k^{\mathrm{in}}=b}^{K} p_{\mathrm{d}}(k^{\mathrm{in}}) k^{\mathrm{in}} \frac{c(k^{\mathrm{in}})}{c_{\mathrm{out}}} \mathcal{B}(\beta_{\mathrm{u}}^{\mathrm{ch}}, k^{\mathrm{in}}, 0) \\ & + \zeta \sum_{k^{\mathrm{in}}=K+1}^{\infty} p_{\mathrm{d}}(k^{\mathrm{in}}) \sqrt{k^{\mathrm{in}}} \frac{c(k^{\mathrm{in}})}{c_{\mathrm{out}}}. \end{aligned} \tag{4.90}$$

The convergence of $\beta_{\mathrm{u}}^{\mathrm{ch}}$ and $\beta_{\mathrm{l}}^{\mathrm{ch}}$ to their asymptotic value β^{ch} in function of K has been plotted in figure 4.4. Since $\beta_{\mathrm{u}}^{\mathrm{ch}}$ saturates much faster we have used this value in figure 4.5 to estimate β^{ch}. The R phase contains a part with $d > 0$ which is chaotic. The bounds $\beta_{\mathrm{u}}^{\mathrm{m}}$ and $\beta_{\mathrm{l}}^{\mathrm{m}}$ on the inverse critical temperature β^{m} of the transition from the chaotic to the non-chaotic R phase are calculated by substituting respectively $m_{\mathrm{l}}^{\mathrm{a}}$ and $m_{\mathrm{u}}^{\mathrm{a}}$ into equation (4.84). In figure 4.5 the value of $\beta_{\mathrm{l}}^{\mathrm{m}}$ is approximated by β^{m}. In this figure the complete phase diagram of the neural network with the P phase, the R phase and the CH phase is presented. The chaotic region of the neural network is enclosed by the dashed lines. This region is larger for odd values of p.

4.10 The stationary solution without cycles

In this section we develop an algorithm to calculate the marginals of the stationary solution. This algorithm would be the equivalent of the BP equations (2.88). When we use equations (4.13) we have an algorithm with a linear computational complexity $\mathcal{O}(N)$. Because the algorithm scales as $\mathcal{O}(2^t)$ in time t we can only follow the dynamics for a small number of time steps. We are interested in calculating the marginals of the stationary state when $t \to \infty$. To solve this problem we introduce some assumptions on the distribution $P(\sigma^{0..t})$. The calculation is restricted to stationary states without cycles.

4.10.1 The one-time approximation

We neglect the correlations in time in the probability distributions $P^{(\ell)}_{i,0..t}$ equations (4.13)

$$P^{(\ell)}_{i,0..t}(\sigma^{0..t}|\theta^{0..t}) = \prod_{s=1}^{t} P_i^{(\ell),*}(\sigma(s)|\theta(s)) = \prod_{s=1}^{t} \frac{\exp\left[\beta u_i^{(\ell)}(\theta(s))\sigma(s)\right]}{2\cosh\left[\beta u_i^{(\ell)}(\theta(s))\right]}. \quad (4.91)$$

We close the equations (4.13) using the ansatz (4.91) in the left hand side and

$$P_i^{(\ell),*}(\sigma(t)|\theta) = \lim_{s\to-\infty} \sum_{\sigma^{s..t-1}} P_i^{(\ell)}(\sigma^{s..t}|\theta^{s+1..t}), \quad (4.92)$$

with $\theta^{s..t}$ the constant vector with components θ, i.e. $\theta^{s..t} = (\theta, \theta, \ldots, \theta)$. We have closed the equations in a way that we recover the exact stationary solution for both fully symmetric and fully asymmetric graphs. The approximation given by (4.91) does not take into account the appearance of cycles. Since we neglected correlations in time we can not expect a good description of the spin glass phase.

4.10.2 Cavity equations in the one-time approximation

Inserting (4.91) and (4.92) in (4.13) we obtain the following MPA

$$u_{j\to i}(\theta)=\frac{1}{2\beta}\sum_{\sigma}\sigma\log\left\{\sum_{\tau_1,\cdots,\tau_{k_j^{\rm in}-1}}\sum_{\tau}\frac{\exp\left[\beta\sigma(\sum_{\ell\in\partial_j^{\rm in}\setminus i}J_{\ell j}\tau_\ell+\theta)\right]}{2\cosh\left[\beta\left(\sum_{\ell\in\partial_j^{\rm in}\setminus i}J_{\ell j}\tau_\ell+\theta\right)\right]}\right.$$
$$\left.\frac{\exp\left[\beta\sum_{\ell\in\partial_j^{\rm sym}\setminus i}\tau_\ell u_{\ell\to j}\left(J_{\ell j}\tau\right)+\beta\sum_{\ell\in(\partial_j^{\rm in}\setminus\partial_j^{\rm sym})\setminus i}\tau_\ell u_{\ell\to j}\right]}{\prod_{\ell\in\partial_j^{\rm sym}\setminus i}\cosh\left[\beta u_{\ell\to j}(J_{\ell j}\tau)\right]}\exp\left[\beta\tau u_{j\to i}(\theta)\right]\right\}. \tag{4.93}$$

Equation (4.93) is an implicit equation in the message $u_{j\to i}(\theta)$. Using the one-time approximation on equations (4.15) we find an equivalent expression for the real marginals.

For fully asymmetric graphs we find

$$u_{j\to i}=\frac{1}{2\beta}\sum_{\sigma}\sigma\log\left\{\sum_{\tau_1,\cdots,\tau_{k_j^{\rm in}}}\prod_{\ell\in\partial_j^{\rm in}}\frac{\exp\left[\beta\sigma J_{\ell j}\tau_\ell\right]}{2\cosh\left[\beta\left(\sum_\ell J_{\ell j}\tau_\ell\right)\right]}\exp\left[\beta\sum_{\ell=1}^{k_j^{\rm in}}\tau_\ell u_{\ell\to j}\right]\right\}. \tag{4.94}$$

These are the ABP equations (4.16).

For fully symmetric graphs the equations (4.93) admit the solution

$$u_{j\to i}(\theta)=\theta+\beta^{-1}\sum_{\ell\in\partial_j^{\rm sym}\setminus i}\operatorname{atanh}\left[\tanh\left(\beta J_{\ell j}\right)\tanh\left(\beta u_{\ell\to j}\right)\right], \tag{4.95}$$

where we used $u_{\ell\to j}\left(J_{\ell j}\tau\right)=J_{\ell j}\tau+u_{\ell\to j}$. We find again the BP algorithm for symmetric graphs (2.87).

Since we find the known BP algorithms for both fully asymmetric and fully symmetric graphs, we wonder whether equations (4.93) also makes a good prediction of the marginals on partially asymmetric graphs. We determine the performance of the algorithm (4.94) through their DE equations.

4.10.3 Density evolution equations in the one-time approximation

We derive a selfconsistent expression for the distribution of the fields $u_{j\to i}(\theta)$ on a graph drawn from the ensemble $\mathcal{G}_{\rm deg}(N, p_{\rm d})$ from appendix B.2.5. In the sequel we use bimodal distributions of the form (4.51). We have then two types of messages propagating along symmetric links: $u(\theta) = u(J_0\sigma) = u^\sigma$.

The function $u = \mathcal{U}_{k^{\rm in},k^{\rm sym}}(\{u_\ell, J_\ell\}; \theta)$ is defined through the explicit solution u of the following implicit equation:

$$u = \frac{1}{2\beta}\sum_\sigma \sigma \log\left\{\sum_{\boldsymbol{\tau},\tau} \frac{\exp\left[\beta\sigma(\sum_{\ell=1}^{k^{\rm in}} J_\ell\tau_\ell + \theta)\right]}{2\cosh\left[\beta\left(\sum_{\ell=1}^{k^{\rm in}} J_\ell\tau_\ell + \theta\right)\right]}\right.$$
$$\left.\frac{\exp\left[\beta\sum_{\ell=1}^{k^{\rm sym}} \tau_\ell u_\ell\,(J_\ell\tau) + \beta\sum_{\ell=k^{\rm sym}+1}^{k^{\rm in}} \tau_\ell u_\ell\right]}{\prod_{\ell=1}^{k^{\rm sym}} \cosh\left[\beta u_\ell (J_\ell\tau)\right]} \exp\left[\beta\tau u\right]\right\}. \tag{4.96}$$

The density of the fields propagating along directed edges within the one-time approximation is given by

$$W^{\rm d}(u) = \sum_{k^{\rm out}\geq 0}^{\infty}\sum_{k^{\rm in}\geq 0}^{\infty}\sum_{k^{\rm sym}=0}^{\min(k^{\rm out},k^{\rm in})} \frac{p_{\rm d}(k^{\rm in},k^{\rm out},k^{\rm sym})(k^{\rm out}-k^{\rm sym})}{c_{\rm out}-c_{\rm sym}}$$
$$\times\prod_{\ell=1}^{k^{\rm sym}}\int dJ_\ell P_I(J_\ell)\int du_\ell^+ du_\ell^- W^{\rm sym}(u_\ell^+,u_\ell^-)$$
$$\times\prod_{\ell=k^{\rm sym}+1}^{k^{\rm in}}\int dJ_\ell P_I(J_\ell)\int du_\ell W^{\rm d}(u_\ell)$$
$$\times\delta\left[u-\mathcal{U}_{k^{\rm in},k^{\rm sym}}(\{u_{\ell'},J_{\ell'}\}_{\ell'=1..k^{\rm in}};0)\right]. \tag{4.97}$$

The density of the fields propagating along the symmetric edges is

$$W^{\rm sym}(u^+,u^-) = \sum_{k^{\rm in}\geq 0}^{\infty}\sum_{k^{\rm sym}=0}^{k^{\rm in}} \frac{p_{\rm d}(k^{\rm in},k^{\rm sym})k^{\rm sym}}{c_{\rm sym}}$$
$$\times \prod_{\ell=1}^{k^{\rm sym}-1}\int dJ_\ell P_I(J_\ell)\int du_\ell^+ du_\ell^- W^{\rm sym}(u_\ell^+,u_\ell^-)$$
$$\times \prod_{\ell=k^{\rm sym}}^{k_{\rm in-1}}\int dJ_\ell P_I(J_\ell)\int du_\ell W^{\rm d}(u_\ell)$$
$$\times\delta\left[u^- - \mathcal{U}_{k^{\rm in}-1,k^{\rm sym}-1}\left(\{u_{\ell'},J_{\ell'}\}_{\ell'=1..k^{\rm in}-1};-1\right)\right]$$
$$\times\delta\left[u^+ - \mathcal{U}_{k^{\rm in}-1,k^{\rm sym}-1}\left(\{u_{\ell'},J_{\ell'}\}_{\ell'=1..k^{\rm in}-1};1\right)\right]. \qquad (4.98)$$

The fields of the real marginals have a density $W^{\rm real}$ given by

$$W^{\rm real}(u) = \sum_{k^{\rm in}\geq 0}^{\infty}\sum_{u=0}^{k^{\rm in}} p_{\rm d}(k^{\rm in},k^{\rm sym})\prod_{\ell=1}^{k^{\rm sym}}$$
$$\times\int dJ_\ell P_I(J_\ell)\int du_\ell^+ du_\ell^- W^{\rm sym}(u_\ell^+,u_\ell^-)$$
$$\times \prod_{\ell=k^{\rm sym}+1}^{k^{\rm in}}\int dJ_\ell P_I(J_\ell)\int du_\ell W^{\rm d}(u_\ell)$$
$$\times\delta\left[u - \mathcal{U}_{k^{\rm in},k^{\rm sym}}\left(\{u_{\ell'},J_{\ell'}\}_{\ell'=1..k^{\rm in}};0\right)\right]. \qquad (4.99)$$

Using the analogy with the equations (2.35) we call the equations (4.97) and (4.98) the density evolution equations in the one-time approximation. We have now two densities instead of one: the density for the fields propagating along symmetric edges and the density for the fields propagating along directed edges. The equation for the marginals on the original graph (4.99) is the equivalent of (2.37) for the symmetric model. Another important difference is that the update function $\mathcal{U}_{k^{\rm in},k^{\rm sym}}$, for partially asymmetric graphs, is not explicitly known. Instead we have to solve the implicit equation (4.96). The equations (4.97), (4.98) and (4.99) constitute an algorithm that generalizes BP to graphs with asymmetric bonds.

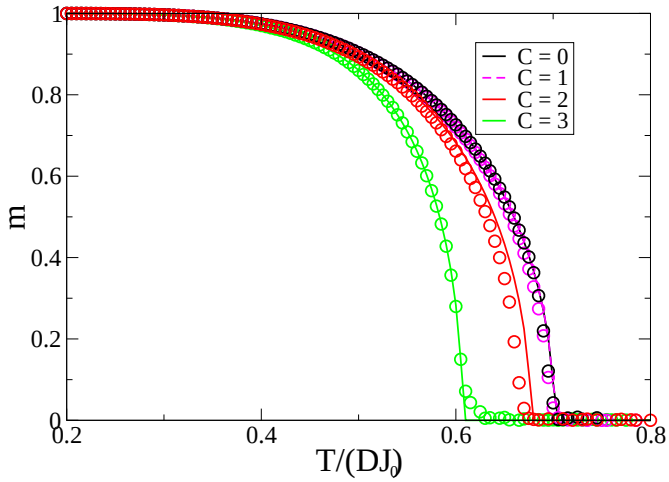


Figure 4.7: The magnetisation m as a function of the rescaled temperature $T/(DJ_0)$ for a ferromagnet (i.e. $\rho = 1$) defined on a Bethe lattice. The indegree D of the graph equals 3 and each site is incident to C symmetric bonds. The simulations (markers) are mean values of 20 runs on graphs of sizes $\mathcal{O}(10^4)$. The theory (lines) follow from solving recursively the density evolution equations in the one-time approximation .

4.10.4 Results and comparison with simulations

We numerically solve the equations (4.97), (4.98) and (4.99) with the population dynamics algorithm as described in subsection 2.2.3. In figure 4.7 the magnetisation is plotted as a function of the temperature for an Ising model without bond disorder on a Bethe lattice. Theory is compared with simulations. The degree distribution is then:

$$p_{\mathrm{d}}(k^{\mathrm{in}}, k^{\mathrm{out}}, k^{\mathrm{sym}}) = \delta(k^{\mathrm{sym}} - C)\delta(k^{\mathrm{in}} - D)p(k^{\mathrm{out}}). \tag{4.100}$$

Since there is no disorder the distributions $W^{\mathrm{d}}(u)$ and $W^{\mathrm{sym}}(u^+, u^-)$ are delta functions. Figure 4.7 shows that the theory and the simulations are in good agreement. For $C = 0, 3$ the results coincide while for $C = 1, 2$ there is a small deviation. Figure 4.8 shows the comparison between theory and simulations on graphs with fluctuating connectivities. We simulated an Ising model without bond disorder on Poissonian graphs drawn from the Poissonian ensemble defined in B.2.4.

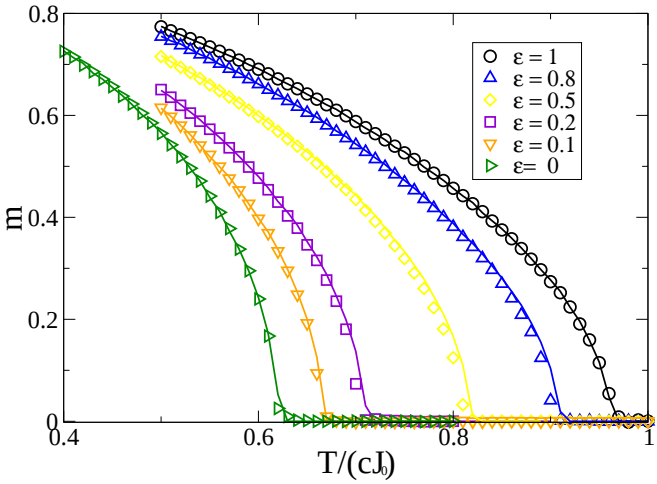


Figure 4.8: The magnetisation m as a function of the temperature $T/(cJ_0)$ for an Ising model without bond disorder on a Poissonian graph with mean connectivity $c = 3$ and different fractions of symmetric edges ϵ. The lines are obtained by population dynamics from the density evolution equations in the one-time approximation for populations of sizes $\mathcal{O}(10^5)$. The markers are the average results from 20 runs with the heat-bath algorithm on a graph instance of size $\mathcal{O}(10^5)$.

Despite the fluctuations in the degrees we find a good agreement between theory and simulations. This confirms that the one-time approximation works very well for models without bond disorder. In figure 4.9 we plot the magnetisation on a Bethe lattice with bond disorder. The difference between theory and simulations increases with the bias ρ in the bonds. From this figure we can already conclude that for partially asymmetric graphs the one-time approximation typically over estimates the value of the magnetization.

4.10.5 Bifurcation analysis

We calculate the P to F and the P to SG transition lines for Ising models with bimodal distributions using a bifurcation analysis around the P solution. First,

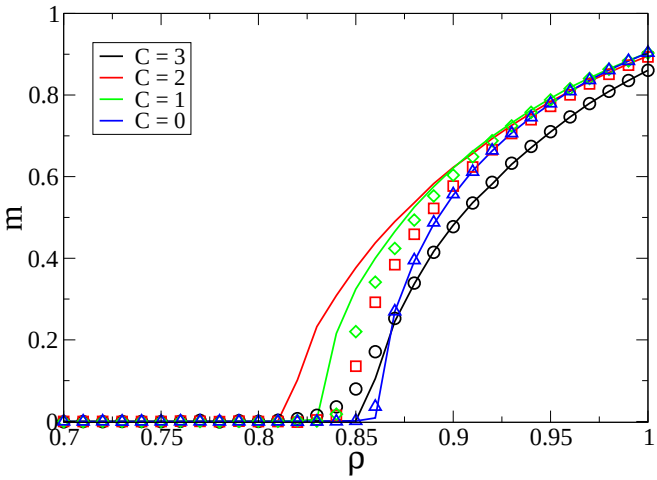


Figure 4.9: The magnetisation m as a function of the bias ρ in the couplings (see equation (4.51)) at a temperature $T/(DJ_0) = 0.5$. The Ising model is defined on a Bethe lattice with the degree distribution defined in (4.100). The indegree D equals 3 and results are shown for various values of the symmetric degree C. The simulations (markers) are the mean values of 20 runs on graphs of sizes $\mathcal{O}(10^5)$. The theory (lines) are results from the density evolution equations in the one-time approximation using a Monte Carlo calculation with populations of $\mathcal{O}(10^5)$ fields.

we note that the equations (4.98) and (4.97) admit the solution:

$$W^{\mathrm{d}}(u) = \delta(u), \tag{4.101}$$

$$W^{\mathrm{s}}(u^+, u^-) = \int dA\, W^P(A)\delta(u^+ - A)\delta(u^- + A). \tag{4.102}$$

Indeed, when we insert (4.101) and (4.102) in (4.98) we get for $W^P(A)$:

$$W^P(A) = \sum_{k^{\mathrm{in}}\geq 0}^{\infty} \sum_{k^{\mathrm{sym}}=0}^{k^{\mathrm{in}}} \frac{p_{\mathrm{d}}(k^{\mathrm{in}}, k^{\mathrm{sym}})k^{\mathrm{sym}}}{c_{\mathrm{sym}}} \prod_{0<\ell\leq k^{\mathrm{sym}}-1} \int dA_\ell W^P(A_\ell)$$

$$\times\delta\left[A - \mathcal{A}_{k^{\mathrm{sym}}-1,k^{\mathrm{in}}-1}\left(\{A_{\ell'}\}_{\ell'=1..k^{\mathrm{sym}}-1}\right)\right], \tag{4.103}$$

with $\mathcal{A}_{k^{\rm sym}-1,k^{\rm in}-1}$ the explicit solution of

$$A = \frac{1}{2\beta}\sum_{\theta=\pm 1}\theta \log\left\{\sum_{\boldsymbol{\tau}}\frac{\exp\left[\theta\beta J_0(\sum_{\ell=1}^{k^{\rm in}-1}\tau_\ell+1)\right]}{\cosh\left[\beta J_0\left(\sum_{\ell=1}^{k^{\rm in}-1}\tau_\ell+1\right)\right]}\cosh\left[\beta\sum_{\ell=1}^{k^{\rm sym}-1}\tau_\ell A_\ell+\beta A\right]\right\}. \tag{4.104}$$

The solution (4.101) and (4.102) represents the P phase. For a symmetric lattice with $k^{\rm in}=k^{\rm sym}$ we have $\mathcal{A}_{k^{\rm sym}-1,k^{\rm in}-1}=J_0$. For $k^{\rm sym}=0$ we have

$$\mathcal{A}_{k^{\rm sym}-1,k^{\rm in}-1} = \frac{1}{2\beta}\sum_{\theta=\pm 1}\theta \log\left\{\sum_{\boldsymbol{\tau}}\frac{\exp\left[\theta\beta J_0(\sum_{\ell=1}^{k^{\rm in}-1}\tau_\ell+1)\right]}{\cosh\left[\beta J_0\left(\sum_{\ell=1}^{k^{\rm in}-1}\tau_\ell+1\right)\right]}\right\}. \tag{4.105}$$

For partially asymmetric graphs $W^P(A)$ will have a more complicated form.

To determine the P to F and P to SG transitions, we expand the equations (4.97), (4.98) and (4.99) around the solution given by (4.101) and (4.102). We refer to [126] for more details. For fully symmetric lattices we get the bifurcation conditions (2.46) and (2.47) back, while for fully asymmetric lattice we find the bifurcation condition (4.57).

4.10.6 Phase diagrams

We determine the phase diagrams of two Ising models on asymmetric graphs within the one-time approximation for a bimodal distribution $P_I(J)$ (4.51). The P-F and P-SG transitions are determined as the solution to the bifurcation conditions found in subsection 4.10.5. The F-SG transitions are determined through a version of the population dynamics algorithm , see subsection 2.2.3, adapted to find the solution to the equations (4.97), (4.98) and (4.99).

- *Bond-disordered Ising model on a Bethe lattice*: In figure 4.9 the $(T/(J_0D),\rho)$-phase diagram of the Ising model on a graph drawn from the ensemble $\mathcal{G}_{\rm Bethe}$ is shown. We see how within the one-time approximation the SG phase gradually disappears with decreasing connectivity in favor of the P and F phase. One should be careful when interpreting this phase diagram. The one-time approximation overestimates the SG phase when there is partial asymmetry.
- *Bond-disordered Ising model on a Poissonian lattice*: In figure 4.10 we present the P to F transition line for a ferromagnet ($\rho=1$) and the P

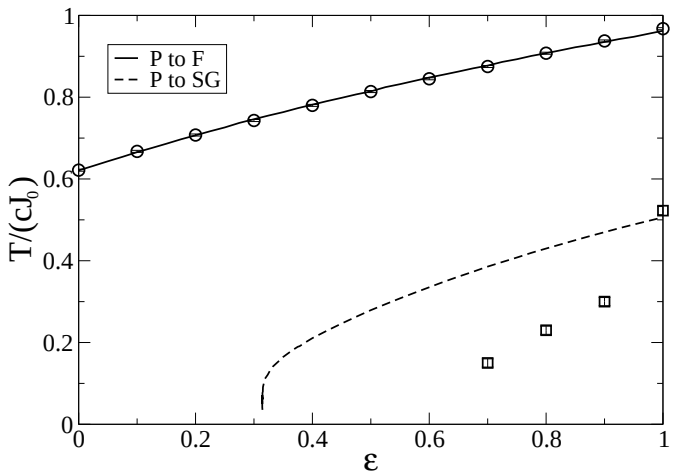


Figure 4.10: Critical temperatures computed from a bifurcation analysis (lines) within the one-time approximation are compared with simulations (markers) as a function of the fraction of symmetric edges ϵ. The results shown for the P to F transition are for an Ising model on a Poissonian graph without bond disorder. The results shown for the P to SG transition concern an Ising model on a Poissonian graph for $\rho = 0$.

to SG transition line for a Poissonian graph as a function of the fraction of symmetric edges ϵ. Both lines are computed with the bifurcation analysis. The markers show the P to F transitions (circles) and the P to SG transitions (squares) both obtained through simulations. The P to F transition is in good agreement with simulations but the P to SG transition does not agree with simulations. The size of the SG phase is overestimated. But qualitatively the one-time approximation does show a SG phase gradually disappearing with the asymmetry in the system. The critical temperatures are obtained through the method of Binder cumulants [15].

4.11 Conclusion

In this chapter we have applied the cavity method to study the dynamics of spin models on a given graph instance. We have derived a set of effective equations which describe the dynamics. Solving these recursive equations can be seen as the

equivalent of the belief propagation algorithm known from inference problems or decoding algorithms. Just like the latter, we expect these equations to be exact on a tree. The main difference with statics is that now path probabilities, instead of stationary probabilities of single spins, are propagated along the edges of the graph. We took the average over an ensemble of graphs to find the recursive equations describing the dynamics of Ising models on typical graphs drawn from this ensemble. These equations generalize the result of [74] to graphs with arbitrary degree distributions. The macroscopic evolution of the system is given as a function of three mean values of path probabilities: one of the probabilities propagating along directed edges, one of the probabilities propagating along symmetric edges and one of the marginal path probabilities of the spins on the original graph. On a Poissonian graph the three recursive equations for these path probabilities reduce to one and we find back the result derived in [74].

We have solved the following problems with these equations:

The evolution of macroscopic observables like the magnetisation is calculated for a small number of time steps because the computational complexity of the final equations scales exponentially in time. These results are compared with simulations and both are in good agreement. This confirms that our approach determines the dynamics of spin models on graphs.

The effective equations simplify on fully asymmetric graphs. The single spin effective dynamics becomes Markovian and the stationary distribution is derived. We have calculated the phase diagrams of an Ising model on a Bethe lattice showing the existence of a paramagnetic, ferromagnetic and a chaotic phase. We also have studied a neural network with Hebbian couplings on a scale-free graph and we have found that correlations between indegrees and outdegrees increase the retrieval properties of the neural network tremendously.

We have introduced an heuristic method to derive the stationary distribution of Ising like models on partially asymmetric graphs by assuming that the path probabilities factorize in time. Hence we do not consider cyclic behaviour. The equations are closed such that we recover the correct stationary distribution for models on fully symmetric and fully asymmetric graphs. This method is approximative. We found a set of equations that generalize belief propagation to an algorithm that computes the marginals of the stationary distribution of Ising models on graphs with asymmetric couplings. The evolution of macroscopic observables and the phase transitions are obtained with this method and compared with simulations. Without bond disorder the theory and simulations give almost similar results. When introducing bond disorder the theoretical predictions start to fail and a more refined approximation is necessary.

4.12 Future prospects

We mention some possible future developments

- *Sequential dynamics*: In [116] the Langevin dynamics of spin models on a Poissonian lattice has been studied using a GFA approach. The resultant order parameter equations concern a distribution of superfields:

$$c\left[\Psi\right] = \frac{1}{N}\sum_{i=1}^{N}\delta\left(\Psi - \Phi_i\right) \tag{4.106}$$

  with $\Phi_i(t,\theta,\overline{\theta})$ some superfield appearing in the generating functional, i.e. :

$$\Phi_i(t,\theta,\overline{\theta}) = s_i(t) + \overline{\theta}\psi_i(t) + \overline{\psi}_i(t)\theta + i\hat{s}_i(t)\theta\overline{\theta}. \tag{4.107}$$

  with θ and $\overline{\theta}$ Grasmann variables. Is it possible to recover the saddle point equations for $c\left[\Psi\right]$ in a probabilistic way using the cavity method? Can one give a probabilistic interpretation for the variable $\Phi_i(t,\theta,\overline{\theta})$ as the field of a single site marginal on the cavity graph or a message on a given graph instance?

- *Loop corrections*: The BP algorithm (2.88) has been generalized to an algorithm containing loop corrections. How can one include loop corrections to the dynamical equations (4.13)? How do these loop corrections apply to the dynamics on a directed graph (4.16) which are Markovian for graphs without loops? This research direction is in particular interesting as many loops will induce directed "currents" in the graph. How can one describe the appearance of these currents in the model?

- *A functional approach*: is it possible to extend the one-time approximation, see subsection (4.10.1), and the MPA (4.93) to an algorithm that determines the marginals, also in the presence of frustration in the lattice? We are adressing this problem through the variational approach as explained, for instance in [176]. Is it possible to generalize the equations (4.93) in a systematic way to including time correlations?

Chapter 5

Low-density parity-check codes

Low-density parity-check (LDPC) codes are an active research area in information theory. These codes are efficient and reach the Shannon limit for symmetric memoryless channels. Reaching the Shannon limit was the fundamental problem of coding theory since their existence.

For asymmetric channels and channels with memory in the noise it is unclear whether the LDPC codes reach the Shannon limit. Asymmetric channels are important in communication through optical fibers while memory effects are important in wireless communication. We determine the typical performance of LDPC codes on these channels by mapping them on spin models on graphs. We derive the density evolution equations for the binary asymmetric channel and the Gilbert-Elliot channel through the replica method. Such a calculation was previously only known for symmetric channels. The Gilbert-Elliot channel is commonly used to model memory effects. For unbiased sources we recover in particular the results of Wang et. al for asymmetric channels and of Eckford et. al for the Gilbert-Elliot channel [172, 44].

We determine the threshold values of the noise for decoding using the belief-propagation algorithm and decoding employing thermodynamic averages. For noise values above these thresholds the corresponding decoding algorithm fails while it is successful below these values. Using methods of FiC-systems we determine the performance of the decoding algorithm when the true noise of the channel is not known. The temperature characterizes this lack of knowledge. We identify phases in the temperature versus noise phase diagram corresponding to spin models on graphs: ferromagnetic, suboptimal ferromagnetic, paramagnetic and spin-glass phases are identified. Only in the ferromagnetic phase reliable communication is possible. In particular we discuss the influence of asymmetry and memory effects on the phase diagram.

5.1 Processes in reliable communication through a noisy medium

We consider the problem of reliable communication between a person A and another person B. Person A sends a message through a noisy channel which corrupts the message. Person B has to infer the message sent from the corrupted message received. We describe very shortly the basic concepts from information and coding theory. For more details we refer to [32, 96].

5.1.1 Source of the information

The source of messages, which we want to communicate reliably, is modelled as a memoryless random number generator that draws strings of bits $\left\{s_i^0\right\}_{i=1..N}$ from a probability distribution $p_{\rm s}$:

$$p_{\rm s}\left(\left\{s_i^0\right\}_{i=1..N}\right) = \prod_{i=1}^{N} p_{\rm s}(s_i^0|b) = \prod_{i=1}^{N} \left[b\delta\left(s_i^0;0\right) + (1-b)\delta\left(s_i^0;1\right)\right]. \tag{5.1}$$

The parameter b is the bias of the source. The index i can be interpreted as a time index. A string of bits $\left\{s_i^0\right\}_{i=1..N} = (s_1^0, s_2^0, \cdots, s_N^0)$ is usually called a word. An important property of the source is the number of information bits contained in each bit sent, given by the binary entropy $h(p_{\rm s})$:

$$h(p_{\rm s}) = \lim_{N\to\infty} -N^{-1} \sum_{\left\{s_i^0\right\}_{i=1..N}} p_{\rm s}\left(\left\{s_i^0\right\}_{i=1..N}\right) \log_2\left(p_{\rm s}\left(\left\{s_i^0\right\}_{i=1..N}\right)\right). \tag{5.2}$$

We also write $H = Nh(p_{\rm s})$. We define the rate of the source $R_0 \equiv h^{-1}(p_{\rm s})$. The rate of the source is an important quantity in data compression [32]:

- **Equipartition theorem**: For $N \to \infty$ there is a typical set of words of size 2^{Nh} containing a probability almost equal to one and all the words in this typical sets are equiprobable: 2^{-Nh}. This results follows directly from the CLT.
- **Data compression**: the average length of codewords in successful data compression is lower bounded by the entropy

5.1.2 Encoding process

Because of the noise present in the channel, we add redundancy to the original word we want to communicate reliably. The encoding process is defined by a map $\mathcal{G}$ from $\mathbb{Z}_2^N$ to $\mathbb{Z}_2^M$ with $N < M$. The rate of this encoding process is given by $R = R_0\,(N/M)$. The image $\mathcal{C} \subset \mathbb{Z}_2^M$ of $\mathcal{G}$ is called the codebook. The elements of the codebook $\mathcal{C}$ are called the codewords.

An important class of codes are the linear codes for which $\mathcal{G}$ is a linear map. These codes are defined through the $M \times N$ generator matrix $\mathbb{G}$:

$$\mathcal{G}(\boldsymbol{s}^0) = \mathbb{G}^T \boldsymbol{s}^0 = \boldsymbol{s}, \tag{5.3}$$

with $\boldsymbol{s}^0$ and $\boldsymbol{s}$ respectively an N-dimensional and M-dimensional column vector. We call the normal form, $\tilde{\mathbb{G}}$ of a generator matrix $\mathbb{G}$, the matrix of the following form,

$$\tilde{\mathbb{G}} = (I_N|\mathbb{P})\,, \tag{5.4}$$

equivalent to $\mathbb{G}$. The quantity I_N is the unity matrix of dimension N and $\mathbb{P}$ the matrix of parity checks. Two generator matrices are equivalent when they generate the same codebooks. It is possible to find the normal form of a generator matrix through a Gauss-Jordan elimination [142]. In practice one performs a series of n column operations:

$$\mathbb{G}^T C_1 C_2 \cdots C_{n-1} C_n C_n^{-1} C_{n-1}^{-1} \cdots C_2^{-1} C_1^{-1} \boldsymbol{s}^0 = \boldsymbol{s}, \tag{5.5}$$

with C_i a $N \times N$ matrix that performs a column operation, permutes two columns or performs a linear combination of two columns.

It is possible to define linear codes through a parity-check matrix $\mathbb{H}$, which is a $(M-N) \times M$-matrix whose rows are all linearly independent. The codebook of parity-check generated matrices is given by

$$\mathcal{C} = \left\{\boldsymbol{s} \in \mathbb{Z}^M | \mathbb{H}\boldsymbol{s} = \mathbf{0}_{1,M}\right\}, \tag{5.6}$$

with $\mathbf{0}_M$ the M-dimensional null vector. We remark that the linear space spanned by the set of $\mathcal{C}$ is N-dimensional when the rows are linearly independent. The generator matrix is found through the relation $\mathbb{G}\mathbb{H}^T = \mathbf{0}_{N,M-N}$, with $\mathbf{0}_{N,M-N}$ an $N \times (M-N)$ dimensional matrix.

5.1.3 Modelling of the channel

The channel will corrupt the message sent $\boldsymbol{s}$. The channel is defined through a probability distribution $p_{\text{chan}}(\boldsymbol{r}|\boldsymbol{s})$ which gives the chance to receive the message $\boldsymbol{r}$ when the message $\boldsymbol{s}$ has been sent. In memoryless channels the bits get flipped independently: $p_{\text{chan}}(\boldsymbol{r}|\boldsymbol{s}) = \prod_{i=1}^{M} p_{\text{chan}}(r_i|s_i)$. Three different types of memoryless channels are frequently studied in the literature:

- The binary erasure channel (BEC): In this case the channel outputs three symbols: $0, 1$ and the erasure $*$:

$$p_{\text{chan}}(r_i|s_i) = (1-\epsilon)\delta\left(r_i; s_i\right) + \epsilon\delta\left(r_i; *\right), \tag{5.7}$$

  with ϵ the chance for an erasure. The erasure channel is probably the most simple example of a noisy channel and thus a good playground for developing error-correcting codes (ECC).

- The binary symmetric channel (BSC): in this binary channel every bit has a chance p to flip

$$p_{\text{chan}}(r_i|s_i) = (1-p)\delta\left(r_i; s_i\right) + p\delta\left(r_i; s_i \oplus 1\right). \tag{5.8}$$

- The binary-input additive white Gaussian noise channel (BIAWGNC): the output is given by $r = s + n$, with n the additive noise drawn from $\mathcal{N}\left(0, \sigma^2\right)$:

$$p_{\text{chan}}\left(r_i|s_i\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(r_i - s_i)^2}{2\sigma^2}\right). \tag{5.9}$$

5.1.4 The decoding process

At last one should define the precise protocol used in the decoding process: the person B has to infer the message sent from the probability distribution $P(\boldsymbol{s}|\mathcal{G}, \boldsymbol{r})$. The decoding process gives an estimate $\hat{\boldsymbol{s}} = (\hat{s}_1, \hat{s}_2, \cdots, \hat{s}_M)$ of the original message. Some possible decoding algorithms are:

- When decoding through the *maximum a posteriori probability*, also called MAP-decoding, one chooses $\hat{\boldsymbol{s}}$ as follows

$$P\left(\hat{\boldsymbol{s}}|\mathcal{G}, \boldsymbol{r}\right) = \max_{\boldsymbol{s}} P(\boldsymbol{s}|\mathcal{G}, \boldsymbol{r}) \tag{5.10}$$

- the *maximizer through posterior marginals* (MPM) is the following decoding process:

$$\hat{s}_i = \text{sign}\left(\sum_{\text{s}_\text{i}} s_i \sum_{\boldsymbol{s}\setminus s_i} P(\boldsymbol{s}|\mathcal{G},\boldsymbol{r})\right). \tag{5.11}$$

  In general it is very hard to calculate the sum (5.11) since it contains 2^{N-1}-terms.

- *Minimum distance decoding*: in this case one chooses the codeword that minimizes the Hamming distance d_H to the received word

$$d_H(\boldsymbol{r},\hat{\boldsymbol{s}}) = \min_{\boldsymbol{s}\in\mathcal{C}} d_H(\boldsymbol{r},\boldsymbol{s}). \tag{5.12}$$

5.1.5 An example: repetition codes

For a r-repetition code the linear map is defined on $\mathbb{Z}_2$ through:

$$\mathbb{G} = \underbrace{(11\ldots1)}_{r-\text{times}}. \tag{5.13}$$

The codebook consists of:

$$\mathcal{C} = \left\{\underbrace{(11\ldots1)}_{r-\text{times}}, \underbrace{(00\ldots0)}_{r-\text{times}}\right\}. \tag{5.14}$$

If person A wants to send the message (110) he/she encodes it into (111111000). Person B decodes the message using a majority rule in every triplet. This is a minimum distance decoding algorithm: person B chooses the codeword in the codebook $\mathcal{C}$ which minimizes the Hamming distance to the received word. For example, the message (110) becomes (111111000). When the noise influences one bit to (101111000) we can still recover the correct message.

For a BSC the probability of error is equal to:

$$p_e = \sum_{n=(r+1)/2}^{r} \binom{n}{r} p^n (1-p)^{r-n}. \tag{5.15}$$

For $r \to \infty$ we use Stirlings approximation:

$$p_e = \int d\rho \exp\left[r\left(\rho \log(\frac{p}{\rho}) + (1-\rho)\log(1-p) - \log(1-\rho) + \rho \log(1-\rho) \right) \right]. \tag{5.16}$$

We see that $p_e \to 0$ for $r \to \infty$. There is a trade-off between the probability of error and the rate of communication r.

5.2 The Shannon limit

In the previous section we have seen an example of an error-correcting code. For most practical codes, reliable communication through forward error-correction is only possible when the rate becomes zero. A landmark result in information theory is Shannon's channel coding theorem which says that one can decode reliable as long as the rate R is smaller then some number C_s.

5.2.1 Some definitions

- The *channel capacity* C is defined as the maximal achievable rate at which reliable communication is possible over a noisy channel. More specific, C is the maximal rate R such that the probability of error $p_\mathrm{e} \to 0$ for $N \to \infty$. The channel capacity is a quantity which depends only on the properties of the channel, and not on the specific encoding and decoding procedures.
- *Random coding*: random coding is an encoding procedure where the 2^N codewords are drawn randomly from the space of codewords $\mathbb{Z}_2^M$.
- The *conditional entropy* of a joint distribution $p(x,y)$, $H_\mathrm{c}(p(x,y),p(y))$, is given as

$$H_\mathrm{c}(p(x,y),p(y)) \equiv -\sum_{x;y} p(x,y)^{\,2} \log\left(p(x|y)\right). \tag{5.17}$$

- The *mutual information* I is

$$I\left(p_\mathrm{I}, p_\mathrm{chan}, p_\mathrm{O}\right) \equiv H\left(p_\mathrm{I}\right) - H_\mathrm{c}\left(p_\mathrm{I} p_\mathrm{chan}, p_\mathrm{O}\right), \tag{5.18}$$

  with $p_\mathrm{I}(x)$ the input distribution to the channel, $p_\mathrm{I}(x)p_\mathrm{chan}(y|x)$ the joint distribution and $p_\mathrm{O}(y)$ the output distribution.

5.2.2 Channel coding theorem

The channel-coding theorem gives the value of the channel capacity [159]:

$$C_{\mathrm{s}} = \lim_{N\to\infty} N^{-1} \max_{p_{\mathrm{I}}(x)} I(p_{\mathrm{I}}, p_{\mathrm{chan}}, p_{\mathrm{O}}). \tag{5.19}$$

A proof of this theorem is given in [32].

We show that the channel capacity is reached by random coding. Important in this proof is the equipartition theorem for random variables:

- The codewords are drawn randomly. When $N \to \infty$, one can say with probability one that the codewords are drawn from a set of size $2^{MH(p_{\mathrm{I}})}$. As the number of codewords equals 2^N, there is a chance $\frac{2^N}{2^{MH(p_{\mathrm{I}})}} = 2^{-M(H(p_{\mathrm{I}})-R)}$ that a randomly drawn word is codeword. The chance that it is not a codeword is thus $1 - 2^{-M(H(p_{\mathrm{I}})-R)}$
- There are $2^{MH_{\mathrm{c}}(p_{\mathrm{I}}p_{\mathrm{chan}},p_{\mathrm{O}})}$ inputs to the channel that correspond with a given output. The chance that one can decode correctly is equal to the chance that $2^{MH_{\mathrm{c}}(p_{\mathrm{I}}p_{\mathrm{chan}},p_{\mathrm{O}})} - 1$ of these words are not codewords:

$$\begin{aligned} p_{\mathrm{correct}} &= \left[1 - 2^{-M[H(p_{\mathrm{I}})-R]}\right]^{2^{MH(p_{\mathrm{I}}p_{\mathrm{chan}},p_{\mathrm{O}})-1}} \\ &\approx 1 - 2^{-M[H(p_{\mathrm{I}})-R-H(p_{\mathrm{I}}p_{\mathrm{chan}},p_{\mathrm{O}})]}. \end{aligned} \tag{5.20}$$

  The chance $p_{\mathrm{correct}} = 1$ as long as $R < C_{\mathrm{s}}$.

5.2.3 Examples

The capacities of the channels defined in subjection 5.1.3 are:

- The capacity of the BEC:

$$C_{\mathrm{BEC}}(\epsilon) = 1 - \epsilon. \tag{5.21}$$

- The capacity of the BSC:

$$C_{\mathrm{BSC}}(p) = 1 - h(p), \tag{5.22}$$

  with $H_b(p) = -p\log_2(p) - (1-p)\log_2(1-p)$.

- The capacity of the BIAWGNC:

$$C_{\text{BIAWGN}}(\sigma) = -\int dx \phi_\sigma(x) \log_2 \phi_\sigma(x) - \frac{1}{2}\log_2\left(2\pi e\sigma^2\right), \tag{5.23}$$

with

$$\phi_\sigma(x) = \frac{1}{\sqrt{8\pi\sigma^2}}\left[\exp\left(-\frac{(x+1)^2}{2\sigma^2}\right) + \exp\left(-\frac{(x-1)^2}{2\sigma^2}\right)\right]. \tag{5.24}$$

5.3 Low-density parity-check codes

Using random codes one can reach the maximum admissible information rate C_{s}. However, random codes are unpractical. One needs a memory of the size 2^N to store all the codewords. Moreover, there is no fast decoding process for random codes [13] (i.e. a decoding process with a computational complexity that scales polynomial in N). The problem of the existence and construction of practical codes reaching the Shannon limit remains. This fundamental problem of coding theory has been solved by the invention of low-density parity-check codes (LDPC). For erasure channels it has been shown that LDPC codes reach the Shannon limit [93] while for BSC and AWGNC it has been shown that they approach closely the Shannon limit [146].

The decoding algorithm of LDPC codes has a linear complexity and the encoding algorithm has a quadratic complexity in the codeword size. It is possible to achieve a linear encoding complexity [148]. As LDPC are linear codes the codewords are stored through the knowledge of a parity-check matrix.

LDPC codes are used in a protocol for digital video broadcasting through satellite transmission [3] and the 802.16e WiMax standard for wireless transmission of data. In these applications high rates are important.

5.3.1 The encoding process

The codewords of linear codes form a linear space and are therefore easy to store. Most practical codes (Hamming codes, repetition codes, Reed-Solomon codes) have a zero rate under reliable communication. To make linear codes that reach the Shannon limit and have a fast decoding algorithm, Gallager suggested linear codes where the parity-check matrix $\mathbb{H}$ is drawn as follows:

- The matrices are random
- The matrices are sparse, i.e. they contain a finite number of ones in each row and column

The randomness of the codes makes them good candidates to reach the Shannon limit. The sparseness makes sure that the decoding algorithm will be efficient.

We have the following type of codes:

- (C, K)-*regular LDPC codes*: the parity-check matrix is a random matrix which has in each column exactly K ones and in each row exactly C ones
- *irregular LDPC codes*: the parity-check matrix is a random matrix for which the number of ones in each row is drawn from a degree distribution c_l and the number of ones in each column is drawn from a degree distribution v_k. The degree distributions are specified by their generating functions:

$$c(x) = \sum_{l=2}^{\infty} c_l x^l, \tag{5.25}$$

$$v(x) = \sum_{k=0}^{\infty} v_k x^k. \tag{5.26}$$

  One obtains regular codes setting $v_k = \delta(k; C)$ and $c_l = \delta(l; K)$.

These LDPC codes are slight variations on the ones introduced by Gallager. Regular codes outperform the usual coding schemes used in most protocols, see [94, 95] or [96]. To obtain codes that reach or approach the Shannon limit (i.e. the rate of the code equals the capacity of the channel), one has to optimize the performance the rate of the code to the degree distributions v_k and c_l.

5.3.2 The posterior probability distribution

To map error-correcting codes on Ising models on graphs we make the transformation $\sigma = (-1)^s$, going from a $(\{0,1\}, \oplus)$-representation to a $(\{-1,1\}, .)$-representation. Addition in modulo 2 becomes a multiplication. We will denote the bits in antipodal signalling using the equivalent Greek letters. We give an example: the parity-check $s_1 \oplus s_2 \oplus s_3 = 0$ becomes $\sigma_1\sigma_2\sigma_3 = 1$. The link between Ising models developed in spin glass theory and error-correcting codes was pointed out by Sourlas [163]. The probability distribution in the antipodal presentation becomes:

$$P_{\mathcal{C}}\left(\{\sigma_i\}_{i=1..M} \,|\, \{\rho_i\}_{i=1..M}\right) \sim \delta_{\mathcal{C}}\left[\boldsymbol{\sigma} \in \mathcal{C}\right] \exp\left(g_i(\rho_i)\sigma_i + f_i\sigma_i\right), \tag{5.27}$$

with the posterior marginals

$$p_{\text{post}}\left(\sigma_i|\rho_i\right) \sim \exp\left(g(\rho_i)\sigma_i + f_i\sigma_i\right). \tag{5.28}$$

5.3.3 The decoding process: belief propagation

Using the sparseness of the parity-check matrix Gallager constructed a decoding algorithm which is a MPA of complexity $\mathcal{O}(M)$. This MPA is best presented on the Tanner graph corresponding with the parity-check matrix $\mathbb{H}$ [166]. The Tanner graph is a bipartite graph of parity-check nodes and variable nodes. Every parity-check node corresponds with a column of the parity-check matrix while every variable node corresponds with a row of the parity-check matrix. A variable node $i = 1..M$ is connected with a parity-check node $a = 1..M - N$ when the corresponding matrix element $\mathbb{H}_{ai} = 1$. When $\mathbb{H}_{ai} = 0$ there is no link between the corresponding nodes. A MPA on the Tanner graph consists of two functions Ψ_{c} and Ψ_v that determine how the messages $u^{(t)}_{a\to i}$ from a parity-check node a to a variable node i and the messages $h^{(t)}_{i\to a}$ from a variable node i to a parity-check node a are updated as a function of messages coming from neighbouring nodes. Given a parity-check function Ψ_{c} and a variable node function Ψ_v the MPA is defined through the following update schemes

$$u^{(t)}_{a\to i} = \Psi_{\mathrm{c}}\left(\left\{h^{(t-1)}_{j\to a}\right\}_{j\in\partial a\setminus i}\right), \tag{5.29}$$

$$h^{(t)}_{i\to a} = \Psi_v\left(\left\{u^{(t)}_{b\to i}\right\}_{b\in\partial i\setminus a}|\rho_i\right), \tag{5.30}$$

and for $t = 0$:

$$u^{(0)}_{a\to i} = 0, \tag{5.31}$$

$$h^{(0)}_{i\to a} = \Psi_v\left(0, 0, \cdots, 0|\rho_i\right). \tag{5.32}$$

The update process (5.29) and (5.30) are shown in, respectively, figures 5.2 and 5.3 for a specific check node and variable node. The total MPA is shown in figure 5.1. The marginals $P_i(\sigma_i)$ are found by:

$$P_i(\sigma_i) = \frac{\exp\left(h_i\sigma_i\right)}{2\cosh\left(h_i\right)}, \tag{5.33}$$

with

$$h_i = \Psi_v\left(\left\{u^{(t)}_{b\to i}\right\}_{b\in\partial i}|g(\rho_i)\right). \tag{5.34}$$

According to the MPM-algorithm defined in subsection 5.1.4, the values of the spins are estimated as

$$\begin{aligned} \hat{\sigma}_i &= \text{sign}\left(\sum_{\sigma} \sigma P_i(\sigma)\right) \\ &= \text{sign}\left(h_i\right). \end{aligned} \tag{5.35}$$

The performance of the algorithm is quantified by the order parameter ρ

$$\rho = M^{-1} \sum_{i=1}^{M} \tau_i \text{sign}\left(\text{h}_\text{i}\right), \tag{5.36}$$

with τ_i the original bits sent . The decoding algorithm given by equations (5.29)-(5.35), has a computational complexity of the order $\mathcal{O}(M)$, which is practically feasible. One can easily decode words of length $\mathcal{O}(1e+6)$ with a present desktop computer. The original problem of taking the sum in (5.11) has a computational complexity $\mathcal{O}(\exp(M))$. The BP algorithm, also called the sum-product algorithm, is given by,

$$\Psi_\text{c}\left(\left\{h_{j\to a}^{(t-1)}\right\}_{j\in\partial a\setminus i}\right) = \text{atanh}\left(\prod_{j\in\partial a\setminus i} \tanh\left(h_{j\to a}^{(t-1)}\right)\right), \tag{5.37}$$

$$\Psi_v\left(\left\{u_{b\to i}^{(t)}\right\}_{b\in\partial i\setminus a}|\rho_i\right) = g(\rho_i) + \sum_{b\in\partial i\setminus a} u_{b\to i}^{(t)}. \tag{5.38}$$

The equations (5.29)-(5.34) are obtained from a cavity calculation on the probability distribution (5.27), see section 2.3.

The message passing algorithm (5.29)-(5.34) is presented graphically in 5.1 on a Tanner graph with 4 check nodes and 6 variable nodes. The message $u_{a\to i}$ propagates from a check node a to a variable node i. The message $h_{i\to a}$ propagates from a variable node i to a check node a. In figures 5.2 and 5.3 we show how, respectively, a message $u_{c\to j}$ and a message $h_{i\to c}$ are updated at time step t.

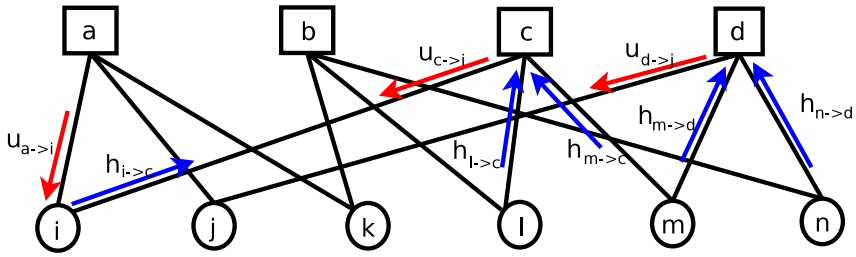


Figure 5.1: We show the message passing algorithm on a Tanner graph corresponding with a $(2, 3)$-regular code. The squares denote check nodes while the circles denote variable nodes.

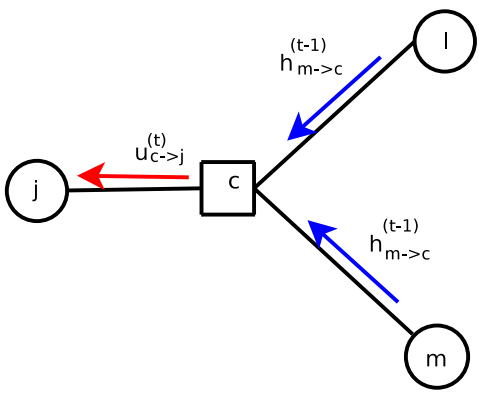


Figure 5.2: We show how the message $u^{(t)}_{c\to j}$ from the check node c to the variable node j is updated with input messages $h^{(t-1)}_{i\to c}$ and $h^{(t-1)}_{m\to c}$.

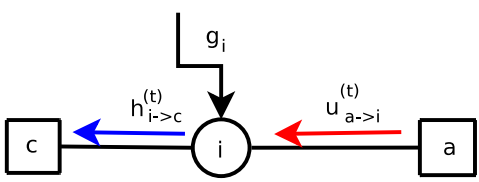


Figure 5.3: We show how the message $h^{(t)}_{i\to c}$ from the variable node i to the check node c is updated with input messages $u^{(t)}_{a\to i}$ and g_i.

5.3.4 Density evolution equations

The DE equations give the ensemble averaged densities W and Z of the messages propagated during the sum-product algorithm:

$$\begin{aligned} W_{\mathrm{c}}^{(t)}(h) &= \sum_{k=0}^{\infty} \frac{v_k k}{\overline{k}} \int \left(\prod_{r=1}^{k-1} du_r Z_{\mathrm{c}}^{(t-1)}(u_r) \right) \\ &\quad \times \sum_{\rho} p(\rho)\, \delta \left[h - g(\rho) - \sum_{r=1}^{k-1} u_r \right], \end{aligned} \tag{5.39}$$

$$\begin{aligned} Z_{\mathrm{c}}^{(t)}(u) &= \sum_{l=2}^{\infty} \frac{c_l l}{\overline{l}} \int \left(\prod_{r=1}^{l-1} dh_r W_{\mathrm{c}}^{(t-1)}(h_r) \right) \\ &\quad \times \delta \left[u - \operatorname{atanh} \prod_{r=1}^{l-1} \tanh(h_r) \right]. \end{aligned} \tag{5.40}$$

The order parameter ρ equation (5.36) is given by

$$\rho = \int dh W(h) \operatorname{sign}(\mathrm{h}), \tag{5.41}$$

with

$$\begin{aligned} W(h) &= \sum_{k=0}^{\infty} v_k \int \left(\prod_{r=1}^{k} du_r Z_{\mathrm{c}}^{(\infty)}(u_r) \right) \\ &\quad \times \sum_{\rho} p(\rho)\, \delta \left[h - g(\rho) - \sum_{r=1}^{k} u_r \right]. \end{aligned} \tag{5.42}$$

The DE equations can be derived from the BP algorithm when the following two conditions are met [147]:

- The graph has a local tree structure

- The system is symmetric, i.e.:

$$p_{\text{chan}}(\rho|\sigma) = p_{\text{chan}}(-\rho| - \sigma) \tag{5.43}$$

$$\Psi_{\text{c}}\left(\sigma_{j_1} h^{(t-1)}_{j_1\to a}, \sigma_{j_2} h^{(t-1)}_{j_2\to a}, \cdots, \sigma_{j_{|\partial a|-1}} h^{(t-1)}_{j_{|\partial a|-1}\to a}\right)$$

$$= \left(\prod_{i=1}^{|\partial a|-1} \sigma_{j_i}\right) \Psi_{\text{c}}\left(h^{(t-1)}_{j_1\to a}, h^{(t-1)}_{j_2\to a}, \cdots, h^{(t-1)}_{j_{|\partial a|-1}\to a}\right), \tag{5.44}$$

$$\Psi_{v}\left(-u^{(t)}_{b_1\to i}, -u^{(t)}_{b_2\to i}, \cdots, -u^{(t)}_{b_{|\partial i|\to i}}| - \rho_i\right)$$

$$= -\Psi_{v}\left(u^{(t)}_{b_1\to i}, u^{(t)}_{b_2\to i}, \cdots, u^{(t)}_{b_{|\partial i|\to i}}|\rho_i\right). \tag{5.45}$$

  When the system is symmetric the error probability is independent of the sent codeword $\{\sigma_i\}_{i=1..M}$. Therefore, one can assume that the all-zero codeword is sent.

One can prove the following concentration theorem [92, 147] on the quantity ρ: For any $\epsilon > 0$ we have

$$P\left(\rho - \mathbb{E}\left[\rho\right] > MC\epsilon/2\right) \leq 2\exp\left(-\delta\epsilon^2 M\right), \tag{5.46}$$

with δ a constant dependend on C, K and the number of iterations t.

5.4 Interplay between physics and error-correcting codes

We use the techniques of FiC systems and spin glasses from chapter 2 to study the posterior probability distribution (5.27).

5.4.1 Mapping of error-correction on Ising models

The posterior probability distribution (5.27) of the codewords is:

$$P_{\mathbb{H},\beta}\left(\{\sigma_i\}_{i=1..M} \,|\, \{\rho_i\}_{i=1..M}\right) \sim \delta_{\mathbb{H}}\left[\{\sigma_i\}_{i=1..M}\right] \exp\left[\beta \sum_{i=1}^{M} \left(g(\rho_i) + f_i\right)\sigma_i\right], \tag{5.47}$$

with $\mathbb{H}$ the parity-check matrix and $g(\rho_i)$ a quenched field. The probability distribution (5.47) contains two factors, one corresponding with the parity checks and another one with the channel noise:

- *The parity-check constraint*: The hard constraint $\delta_{\mathbb{H}}\left[\{\sigma_i\}_{i=1..N}\right]$, which makes sure that $\{\sigma_i\}_{i=1..N}$ is a codeword, can be written as:

$$\delta_{\mathbb{H}}\left[\{\sigma_i\}_{i=1..M}\right] = \lim_{\gamma\to\infty} \exp\left[\gamma \sum_{\langle j_1,j_2,\cdots,j_K\rangle} \mathcal{T}_{\langle j_1,j_2,\cdots,j_K\rangle}\left(J_{j_1 j_2\cdots j_K}\sigma_{j_1}\sigma_{j_2}\cdots\sigma_{j_K}-1\right)\right], \tag{5.48}$$

where we used the tensor $\mathcal{T}$ defined by

$$\mathcal{T}_{\langle j_1,j_2,\cdots,j_K\rangle} \equiv \begin{cases} 1 & \text{if} \quad \prod_{l=1}^{K}\mathbb{H}_{ij_l}=1 \text{ for some } \ i\in\{1,\ldots,M-N\} \\ 0 & \text{if} \quad \text{otherwise} \end{cases}. \tag{5.49}$$

For regular LDPC codes this tensor is drawn from the ensemble:

$$P(\mathcal{T}) = \frac{1}{\mathcal{M}} \prod_{\langle j_1,j_2,\cdots,j_K\rangle} \left[C\frac{(K-1)!}{M^{K-1}}\delta\left(\mathcal{T}_{\langle j_1,j_2,\cdots,j_K\rangle}-1\right) + \left(1-C\frac{(K-1)!}{M^{K-1}}\right)\delta(\mathcal{T}_{\langle j_1,j_2,\cdots,j_K\rangle})\right] \times \prod_{l=1}^{M}\delta\left(\sum_{\langle j_1,j_2,\cdots,j_K;l\rangle}\mathcal{T}_{\langle j_1,j_2,\cdots,j_K\rangle}-C\right). \tag{5.50}$$

The quantity $\mathcal{M}$ is the normalization constant $\mathcal{M} = e^{-MC}(\frac{C^C}{C!})^M$. We have used the notation $\langle j_1, j_2, \cdots, j_K\rangle$ to denote the ordered set $j_1 < j_2 < \cdots < j_K$ and the notation $\langle j_1, j_2, \cdots, j_K; l\rangle$ to denote an ordered set $\langle j_1, j_2, \cdots, j_K\rangle$ with one of the indices equal to l. For irregular codes the tensor is drawn from the ensemble:

$$P(\mathcal{T}) = \prod_{k=3}^{\infty} P(\mathcal{T}^{(k)}), \tag{5.51}$$

with

$$P(\mathcal{T}^{(k)}) = \frac{1}{\mathcal{M}^{(k)}} \prod_{\langle j_1, j_2, \cdots, j_k \rangle} \left[\overline{C}_k \frac{(k-1)!}{M^{k-1}} \delta \left(\mathcal{T}^{(k)}_{\langle j_1, j_2, \cdots, j_k \rangle} - 1 \right) + \left(1 - \overline{C}_k \frac{(k-1)!}{M^{k-1}} \right) \delta(\mathcal{T}^{(k)}_{\langle j_1, j_2, \cdots, j_k \rangle}) \right] \times \prod_{l=1}^{M_k} \delta \left(\sum_{j_2, \cdots, j_k; j_1 = l} \mathcal{T}^{(k)}_{\langle j_1, j_2, \cdots, j_k \rangle} - c_l^{(k)} \right). \tag{5.52}$$

The numbers $c_l^{(k)}$ give the number of hyperedges of degree k that are connected to the l-th site. The distributions c_l and v_k in (5.25) and (5.26) as a function of the numbers $c_l^{(k)}$ are

$$v_k = M^{-1} \sum_{j=1}^{M} \delta \left(\sum_{n=2}^{\infty} c_j^{(n)}; k \right), \tag{5.53}$$

and

$$c_l = \ell^{-1} \frac{\sum_{j=1}^{M} c_j^{(\ell)}}{\sum_{j=1}^{M} \sum_{n=2}^{\infty} \frac{c_j^{(n)}}{n}}. \tag{5.54}$$

- *The channel noise*: the fields $g(\rho_i)$ and f_i in (5.47) are defined through equation (5.28). When $\beta = 1$ we get the true posterior marginals.

  The fields $g(\rho_i)$, for the channels in subsection 5.1.3, are

  – BEC:

  $$g(\rho_i) = \delta(g - \rho_i \infty) \delta(\rho_i; \pm 1). \tag{5.55}$$

  – BSC:

  $$g(\rho_i) = \rho_i \, \mathrm{atanh}(1 - 2p). \tag{5.56}$$

  – BIAWGNC:

  $$g(\rho_i) = \frac{\rho_i}{\sigma^2}. \tag{5.57}$$

  The fields f_i are given by:

  $$f_i = \begin{cases} \frac{1}{2} \log \left(\frac{b}{1-b} \right), & \text{for} \quad i = 1..N \\ 0, & \text{for} \quad i = N+1..M \end{cases}. \tag{5.58}$$

We have used that the encoding, see subsection 5.3.1, makes the parity-check bits unbiased since the generator matrix $\mathbb{G}$ is dense.

In the language of spin models we would say that the problem of LDPC is a random field Ising model on a hypergraph with infinitely strong hyperinteractions. We expect phase transitions that are qualitatively similar to the phase transitions appearing in p-spin models, see subsection 2.4.2. However, this is not completely true. The hard constraints equation (5.48) will change the phase diagram qualitatively. The phase diagram of LDPC contains features of the REM and of the p-spin model: the phase diagram of LDPC contains an entropy crisis analogous to the one in the REM, see subsection 2.4.4. The phase diagram contains also metastable states which are found in p-spin models [16, 58].

5.4.2 Decoding algorithms from statistical physics

Finite temperatures appear naturally when we map the posterior probability distribution on Ising models (5.47). The temperature is a parameter that quantifies the incorrect knowledge of the noise p or σ in the channel. As an example we consider the BSC. We denote the estimated noise by p^* and the true noise by p:

$$\beta = \frac{\operatorname{atanh}(1-2p^*)}{\operatorname{atanh}(1-2p)}. \tag{5.59}$$

When β is small the estimation of p^* is larger than p and when β is large the estimation of p^* is smaller than p.

Using statistical physics one can extend the decoding processes 5.1.4. We define two new decoding processes:

- The first decoding process is a straightforward extension of the BP algorithm (5.29)-(5.34) and (5.37)-(5.38) to finite temperatures
- We define the thermodynamic decoding process through:

$$\hat{\sigma}_i = \operatorname{sign} \sum_{\sigma_i} \sum_{\{\sigma_j\}_{j=1..M} \setminus \sigma_i} P_{\mathbb{H},\beta}\left(\{\sigma_j\}_{i=1..M} \,|\, \{\rho_j\}_{j=1..M}\right). \tag{5.60}$$

When $\beta = 1$ thermodynamic decoding corresponds with MPM and when $\beta = \infty$ it corresponds with MAP. Thermodynamic decoding is therefore a "natural" extension of the decoding processes as defined in subsection 5.1.4.

In the information theoretic literature finite temperature decoding is in general not considered.

5.4.3 The zero-codeword ansatz

To study finite temperature decoding, we calculate the disorder-averaged free energy $\overline{f}$ for unbiased sources

$$-\beta\overline{f} = \sum_{\mathbb{H}} P(\mathbb{H}) \sum_{\{\tau_i\}_{i=1..M}} \delta_{\mathbb{H}}[\{\tau_i\}_{i=1..M}] \sum_{\{\rho_i\}_{i=1..M}} \prod_{i=1}^{M} p_{\text{chan}}(\rho_i|\tau_i)$$
$$\times \log\left\{ \sum_{\{\sigma_i\}_{i=1..M}} \delta_{\mathbb{H}}[\{\sigma_i\}_{i=1..M}] \exp\left[\beta \sum_{i=1}^{M} g(\rho_i)\sigma_i\right]\right\}, \tag{5.61}$$

with $\{\tau_i\}_{i=1..M}$ the transmitted code words. For symmetric channels where $p_{\text{chan}}(\rho_i\sigma_i|\tau_i\sigma_i) = p_{\text{chdynamicsan}}(\rho_i|\tau_i)$, we can write

$$-\beta\overline{f} = \sum_{\mathbb{H}} P(\mathbb{H}) \sum_{\{\rho_i\}_{i=1..M}} \prod_{i=1}^{M} p_{\text{chan}}(\rho_i|1)$$
$$\times \log\left\{ \sum_{\{\sigma_i\}_{i=1..M}} \delta_{\mathbb{H}}[\{\sigma_i\}_{i=1..M}] \exp\left[\beta \sum_{i=1}^{M} g(\rho_i)\sigma_i\right]\right\}. \tag{5.62}$$

Equation (5.62) gives the value of the free energy when we have sent the all-zero codeword. Therefore, we call this the *zero-codeword ansatz*. As we will see later, the zero-codeword ansatz is not valid anymore for asymmetric channels. This will be the main difficulty in dealing with asymmetric channels.

5.4.4 Summary of results for symmetric channels

A statistical mechanics calculation of the disorder-averaged free energy $\overline{f}$ (5.62) using the replica method/cavity method, leads to the following results [170, 124, 171, 114, 56, 165]

- One recovers the DE equations (5.39) and (5.40) using a RS approximation.

- One recovers the MAP threshold through a calculation of the entropy [115]. The result reads

$$s = -\lim_{N\to\infty} N^{-1} \sum_{\{\sigma_i\}_{i=1..N}} P_{\text{post}}\left(\{\sigma_i\}_{i=1..N} \,|\, \{\rho_i\}_{i=1..N}\right) P\left(\{\rho_i\}_{i=1..N}\right) \times \log\left(P_{\text{post}}\left(\{\sigma_i\}_{i=1..N} \,|\, \{\rho_i\}_{i=1..N}\right)\right). \tag{5.63}$$

  When $s > 0$ MAP decoding is impossible. When $s = 0$ MAP decoding is possible.

- One determines the finite-temperature phase diagram containing a region where reliable communication is possible, the F region, and a region where the corruption of the signal by the channel is too much to decode errorless. The main picture is presented in figure 5.4. For noise levels $p < p_{\text{d}}$ the F-phase is the only phase. When $p > p_{\text{d}}$ a suboptimal (S) phase appears to which the initial state P converges. This state will corrupt the reliable communication through the BP algorithm. However, since $f_{\text{F}} < f_{\text{S}}$ thermodynamic decoding is still possible. At a even higher noise level $p > p_{\text{c}} > p_{\text{d}}$ errors appear also in thermodynamic decoding since f_{F} is not anymore the minimal value of the free energy. All noise levels are smaller than the Shannon limit which is consistent with Shannons' channel-coding theorem.

- We call $\beta = 1$ the Nishimori temperature analogous to the corresponding temperature in spin glasses [77]. At this temperature one can prove rigorously a series of (in)equalities [114] in the energy, free energy, specific heat and other thermodynamic quantities [134]. Moreover, one can prove that the thermodynamic behaviour of the model contains no RSB-effects at the Nishimori temperature.

A 1RSB calculation on symmetric channels has been done in [108] and the error-exponents have been calculated in [162, 121]. In this thesis we extend the statistical physics calculations to channels with memory and asymmetry. These results extend on the known results in information theory.

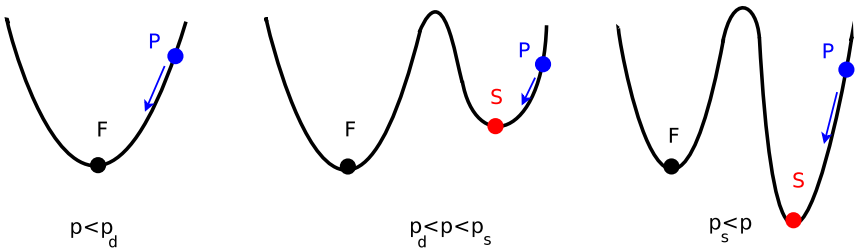


Figure 5.4: The free energy landscape for three different values of the noise p. The initial state of error-correcting algorithms is given by the paramagnetic state P. For $p < p_{\rm d}$ only a F phase is present corresponding with successful decoding. The P state converges to the F-state under iteration of the algorithm. For $p_{\rm d} < p < p_{\rm s}$ a suboptimal ferromagnetic state (S) appears. The P-state lies in the basis of attraction of the S-phase. The S state disrupts the decoding process with the sum-product algorithm. The free energy of the F state is lower than the free energy of any other state in the free energy landscape. An extensive search, through MAP decoding allows us therefore to decode successfully. Above the critical noise level $p_{\rm s}$ MAP decoding fails also.

5.5 LDPC on binary asymmetric channels

5.5.1 Binary asymmetric channels

Definitions

Binary asymmetric memoryless channels (BAC) are simple models for channels with asymmetry where the noise in the channel depends on the value of the bit sent. The BAC is defined through the following single bit transition probability

$$\begin{aligned} p_{\rm chan}(\rho|\sigma) = {} & \delta(\sigma;1)\left[(1-q)\delta(\rho;\sigma) + q\delta(\rho;-\sigma)\right] \\ & +\delta(\sigma;-1)\left[(1-p)\delta(\rho;\sigma) + p\delta(\rho;-\sigma)\right]. \end{aligned} \tag{5.64}$$

The transition probability (5.64) is presented graphically in figure 5.5. We choose in the sequel $p < q$ and write $p = \kappa q$ with $0 \leq \kappa \leq 1$. When $\kappa = 0$ the BAC becomes a Z-channel. The Z-channel is a very simplified model for optical communication; light bits transmitted through an optical fiber can be erased but not created.

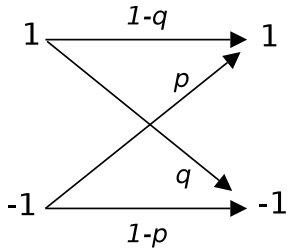


Figure 5.5: A graphical representation of the binary asymmetric channel: noise corrupts the different bits with a different probability.

Mutual information

Introduction of a bias b in the message sent, see (5.1), can increase the mutual information I when communicating the message through an asymmetric channel. The value of the capacity $C_{\rm BAC}(p,q)$ of an asymmetric channel with noises p and q is given by:

$$C_{\rm BAC}(p,q) = I(p,q;b_\star) = h\left(b_\star(1-q)+(1-b_\star)p\right) - \left(b_\star h(q)+(1-b_\star)h(p)\right), \tag{5.65}$$

with the bias $b_\star$ maximizing the mutual information $I(p,q;b)$ of a biased source sending messages through a BAC with noises p and q:

$$b_\star(p,q) = \frac{p\left(\exp\left[F(p,q)\right]+1\right)-1}{(q+p-1)\left(1+\exp\left[F(p,q)\right]\right)}, \text{with} \quad F(p,q) = \left[\frac{h(p)-h(q)}{q+p-1}\right]\log 2. \tag{5.66}$$

As the parity check bits are unbiased the mutual information $I_{\rm G}$ for regular LDPC codes

$$I_{\rm G}\left(\kappa\, q, q; b, \frac{C}{K}\right) = \left(1-\frac{C}{K}\right) I(\kappa\, q, q; b) + \frac{C}{K} I(\kappa\, q, q; 1/2). \tag{5.67}$$

The critical noise levels $q_{\rm s}$ following from $I_{\rm G}$ and C are given by

$$I_{\rm G}\left(\kappa\, q_{\rm s}, q_{\rm s}; b, \frac{C}{K}\right) = R, \tag{5.68}$$

$$C(\kappa q_{\rm s}, q_{\rm s}) = R. \tag{5.69}$$

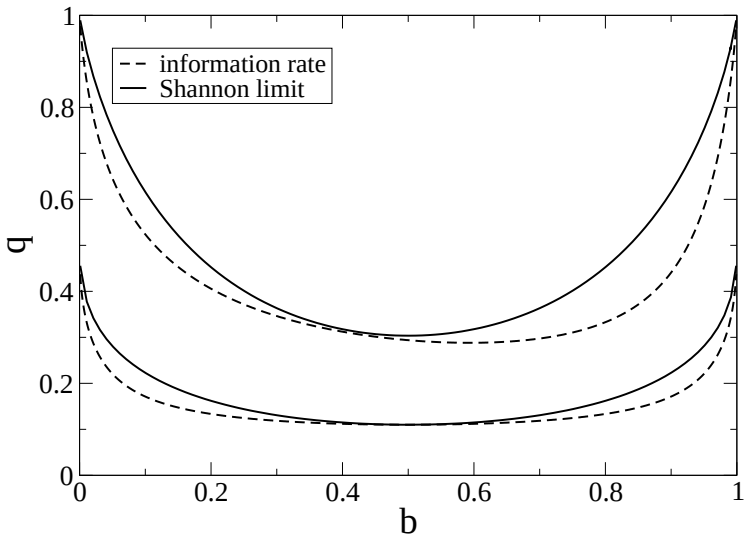


Figure 5.6: The critical noise level q of a BSC (bottom) or Z-channel (top). Using the analogy with LDPC codes we considered a source producing messages at a rate $R = (1 - C/K)h(b)$, with $C/K = 1/2$. The solid lines show the Shannon limit for this rate. The dashed lines show the limit following from the information rate $I_{\rm G}$ of a non-uniform source. This means a source where the first $(1 - C/K)M$ bits are biased with bias b, and the last $(C/K)M$ bits are unbiased.

for a given value of the rate R and the bias b. In figure 5.6 we show these values for $q_{\rm s}$ for a BSC and a Z channel. In figure (5.7) we show the value of the bias which maximizes the mutual information of a Gallager code at a given rate. From figures 5.6-5.7 we conclude that bias in the source is important in asymmetric channels. One reaches the Shannon limit at a $b \neq 0.5$ for asymmetric channels. Regular Gallager codes can, however, not reach this limit since the parity check bits are unbiased.

DE equations

A straightforward generalization of the DE equations (5.40) to asymmetric channels is not possible. For asymmetric channels the symmetry condition, see

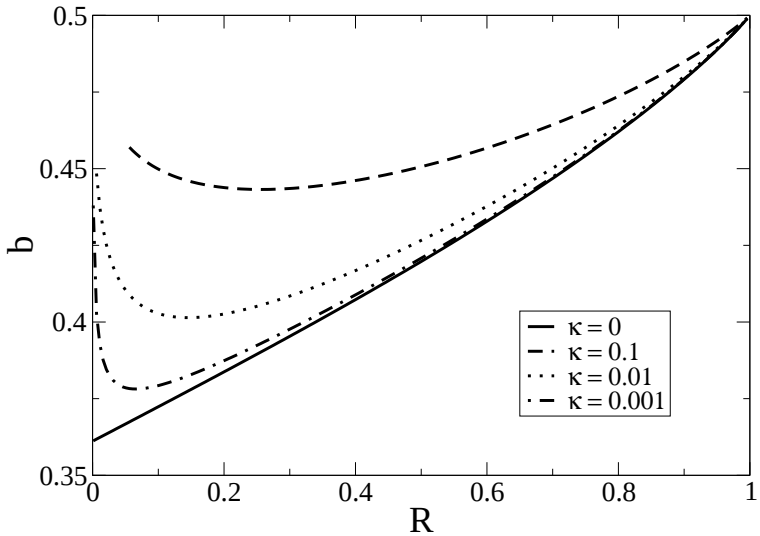


Figure 5.7: The bias b, that maximizes the noise level q following from the mutual information rate of an asymmetric channel coupled with a biased source and a Gallager encoding scheme, as a function of the rate R is shown for several values of the degree of symmetry κ.

subsection 5.3.4, is not fulfilled. The probability distribution is:

$$P_{\mathbb{H},\beta}\left(\{\sigma_i\}_{i=1..N} \,|\, \{\rho_i\}_{i=1..N}\right) \sim \delta_{\mathbb{H}}\left[\{\sigma_i\}_{i=1..N}\right] \exp\left[\beta \sum_{i=1}^{M} \left(g\rho_i + f_i\right) \sigma_i\right] \quad (5.70)$$

with in this case

$$g\left(\rho_i\right) \;=\; \frac{1}{4}\log\left(\frac{(1-p)(1-q)}{pq}\right)\rho_i + \frac{1}{4}\log\left(\frac{(1-q)q}{(1-p)p}\right), \quad (5.71)$$

and f_i still given by (5.58). For $b = 0.5$ we can not assume the zero-codeword ansatz when $p \neq q$. As a consequence the densities of the messages sent in the BP equations will depend on the codeword sent. In [172] the DE equations for asymmetric channels and unbiased sources have been found using the dependence of the densities on the codewords sent.

5.5.2 The Nishimori temperature for asymmetric channels

The disorder-averaged free energy can be written as

$$\overline{f} = \lim_{M\to\infty} \left\langle \sum_{\{\tau_i\}_{i=1..M}} \tilde{P}\left(\{\tau_i\}_{i=1..M}\right) f\left(\{\tau_i\}_{i=1..M}\right) \right\rangle_{\mathbb{H}}, \tag{5.72}$$

with

$$-\beta f\left(\{\tau_i\}\right) = \log \left\{ \sum_{\{\sigma_i\}_{i=1..M}} \delta_{\mathbb{H}}\left[\{\sigma_i\}_{i=1..M}\right] \exp\left[\beta \sum_{i=1}^{M} \left(g(\rho_i) + f_i\right)\sigma_i\right] \right\},$$

and

$$\tilde{P}\left(\{\tau_i\}_{i=1..M}\right) = \frac{\delta_{\mathbb{H}}\left[\{\tau_i\}_{i=1..M}\right] \prod_{i=1}^{N} p_b\left(\tau_i\right) \prod_{i=N+1}^{M} p_{\frac{1}{2}}\left(\tau_i\right)}{\sum_{\{\tau'_i\}_{i=1..M}} \delta_{\mathbb{H}}\left[\{\tau'_i\}_{i=1..M}\right] \prod_{i=1}^{N} p_b\left(\tau'_i\right) \prod_{i=N+1}^{M} p_{\frac{1}{2}}\left(\tau'_i\right)} \tag{5.73}$$

One can show, using the specific form of (5.72), that the following properties hold at $\beta = 1$ [114, 129]:

- The energy $e = \partial_\beta \beta f$ becomes:

$$e(\beta=1) = -\sum_{\{\tau_i\}_{i=1..M}} \tilde{P}\left(\{\tau_i\}_{i=1..M}\right) \times \left[\left(1-\frac{C}{K}\right)\left\langle g(\rho) + f(b)\right\rangle_{\rho|\tau_1} + \frac{C}{K}\left\langle g(\rho) + f\left(\frac{1}{2}\right)\right\rangle_{\rho|\tau_{N+1}}\right]. \tag{5.74}$$

- The system is replica symmetric, i.e. $P^{(2)} = P^{(1)}$, with

$$P^{(1)}(m) = \overline{\langle \delta(m - N^{-1}\sum_i \sigma_i)\rangle_\beta}, \tag{5.75}$$

and

$$P^{(2)}(q) = \overline{\langle \delta(q - N^{-1}\sum_i \sigma_i\sigma'_i)\rangle_\beta}. \tag{5.76}$$

- One can prove $\rho_{\beta=1} \geq \rho_\beta$ for the order parameter (5.36) [133]. This means that thermodynamic decoding performs best when the true noise value is known.

5.5.3 Replica method

Replica trick

We use the replica method as presented in subsection 2.2. The free energy is of the form

$$\overline{f} = \lim_{M\to\infty} \left\langle \sum_{\{\rho_i\}_{i=1..M}} \sum_{\{\tau_i\}_{i=1..M}} \prod_{i=1}^{M} p_{\text{chan}}\left(\rho_i|\tau_i\right) \tilde{P}\left(\{\tau_i\}_{i=1..M}\right) f\left(\{\tau_i\}_{i=1..M}\right) \right\rangle_{\mathbb{H}} \tag{5.77}$$

with

$$-\beta f\left(\{\rho_i\}_{i=1..M}\right) = \log\left\{ \sum_{\{\sigma_i\}_{i=1..M}} \delta_{\mathbb{H}}\left[\{\sigma_i\}_{i=1..M}\right] \exp\left[\beta \sum_{i=1}^{M} \left(g(\rho_i) + f_i\right)\sigma_i\right] \right\}. \tag{5.78}$$

The expression (5.77), describing an average over parity-check matrices and input codewords, can be dealt with using the replica method [107]. We replicate the τ-variables g times obtaining $\boldsymbol{\tau} = \left(\tau^1, \cdots, \tau^g\right)$ and the σ variables n times getting $\boldsymbol{\sigma} = \left(\sigma^1, \cdots, \sigma^n\right)$. The $\boldsymbol{\tau}$ variables are replicated such that the denominator of $\tilde{P}$ gets in the denominator. This is done using the trick

$$\sum_{\tau} \frac{p(\tau)f(\tau)}{\sum_\tau p(\tau)} = \lim_{g\to 0} \sum_{\tau} p(\tau)f(\tau) \left(\sum_{\tau} p(\tau)\right)^{g-1} \tag{5.79}$$

We replicated the summation an integer times and make an analytic continuation to zero. The free energy per bit (5.77) is then given by an extremization problem:

$$-\beta\overline{f} \quad = \quad \lim_{g\to 0} \lim_{n\to 0} \frac{1}{n} \text{extr}_{P,\hat{P}} \Psi\Big\{P(\boldsymbol{\sigma},\boldsymbol{\tau}), \hat{P}(\boldsymbol{\sigma},\boldsymbol{\tau})\Big\}. \tag{5.80}$$

The function Ψ equals

$$\Psi\left\{P(\boldsymbol{\sigma},\boldsymbol{\tau}),\hat{P}(\boldsymbol{\sigma},\boldsymbol{\tau})\right\} = -C\sum_{\boldsymbol{\sigma},\boldsymbol{\tau}}\hat{P}(\boldsymbol{\sigma},\boldsymbol{\tau})P(\boldsymbol{\sigma},\boldsymbol{\tau}) + C - \frac{C}{K}$$
$$+\frac{C}{K}\sum_{(\boldsymbol{\sigma}_1,\boldsymbol{\tau}_1),\cdots,(\boldsymbol{\sigma}_K,\boldsymbol{\tau}_K)}\prod_{l=1}^{K}P(\boldsymbol{\sigma}_l,\boldsymbol{\tau}_l)\prod_{\alpha=1}^{n}\delta\left(\prod_{l=1}^{K}\sigma_l^{\alpha};1\right)\prod_{\zeta=1}^{g}\delta\left(\prod_{l=1}^{K}\tau_l^{\zeta};1\right)$$
$$+\int da q_b(a)\log\left[\sum_{\boldsymbol{\sigma},\boldsymbol{\tau}}\prod_{\zeta=1}^{g}p_a(\tau^{\zeta})\left\langle\hat{P}^C(\boldsymbol{\sigma},\boldsymbol{\tau})\exp\left[\beta\left(g(\rho)+f(a)\right)\sum_{\alpha}\sigma^{\alpha}\right]\right\rangle_{\rho|\tau^1}\right], \tag{5.81}$$

with

$$q_b(a) = \left(1-\frac{C}{K}\right)\delta(a-b)+\frac{C}{K}\delta(a-\frac{1}{2}), \tag{5.82}$$

$$f(a) = \frac{1}{2}\log\left(\frac{a}{1-a}\right), \tag{5.83}$$

and $\langle\cdot\rangle_{\rho|\tau^1}$ an average over $p_{\rm chan}\left(\rho|\tau\right)$. The order parameters $P\left(\boldsymbol{\sigma},\boldsymbol{\tau}\right)$ and $\hat{P}\left(\boldsymbol{\sigma},\boldsymbol{\tau}\right)$ are solutions to the self-consistent equations

$$\hat{P}(\boldsymbol{\sigma},\boldsymbol{\tau}) = \sum_{(\boldsymbol{\sigma}_1,\boldsymbol{\tau}_1),\cdots,(\boldsymbol{\sigma}_{K-1},\boldsymbol{\tau}_{K-1})}\prod_{l=1}^{K-1}P(\boldsymbol{\sigma}_l,\boldsymbol{\tau}_l)$$
$$\times\prod_{\alpha=1}^{n}\delta\left(\sigma^{\alpha}\prod_{l=1}^{K-1}\sigma_l^{\alpha};1\right)\prod_{\zeta=1}^{g}\delta\left(\tau^{\zeta}\prod_{l=1}^{K-1}\tau_l^{\zeta};1\right), \tag{5.84}$$

$$P(\boldsymbol{\sigma},\boldsymbol{\tau}) \sim \int da\ q_b(a)\prod_{\zeta}^{g}p_a(\tau^{\zeta})\left\langle\hat{P}^{C-1}(\boldsymbol{\sigma},\boldsymbol{\tau})\exp\left[\beta\left(g(\rho)+f(a)\right)\sum_{\alpha}\sigma^{\alpha}\right]\right\rangle_{\rho|\tau^1}. \tag{5.85}$$

Inserting (5.84) into (5.85) produces a single self-consistent equation in terms of $P(\boldsymbol{\sigma},\boldsymbol{\tau})$.

Density evolution for asymmetric channels

We write the joint distribution of the replicated spin variables $P(\boldsymbol{\sigma}, \boldsymbol{\tau})$ as:

$$P(\boldsymbol{\sigma}, \boldsymbol{\tau}) = \int da\ q_b(a)\ P(\boldsymbol{\tau}|a)\ P(\boldsymbol{\sigma}|\boldsymbol{\tau}, a). \tag{5.86}$$

In order to take the limit $n \to 0$ in (5.80), one has to make an assumption about the form of the distribution $P(\boldsymbol{\sigma}|\boldsymbol{\tau}, a)$. The simplest such ansatz corresponds to RS, i.e. assuming that the α-replica indices with respect to the noise variables are interchangeable. More concretely we write

$$\begin{aligned} P(\boldsymbol{\sigma}|\boldsymbol{\tau}, a) &= \int dh\, W_{\rm c}(h|\boldsymbol{\tau}, a) \prod_{\alpha=1}^{n} \mathcal{Q}(\sigma_\alpha|h), \\ \mathcal{Q}(\sigma|h) &= \frac{\exp(\beta h \sigma)}{2\cosh(\beta h)}, \end{aligned} \tag{5.87}$$

for some $W_{\rm c}(h|\boldsymbol{\tau}, a)$ with $\int dh\, W_{\rm c}(h|\boldsymbol{\tau}, a) = 1$. Using this ansatz we can convert the self-consistent equation of $P(\boldsymbol{\sigma}, \boldsymbol{\tau})$ into one for the density $W_{\rm c}(h|\boldsymbol{\tau}, a)$ and one for $P(\boldsymbol{\tau}|a)$. The result reads

$$\begin{aligned} P(\boldsymbol{\tau}|a) &\sim \prod_{\zeta} p_a(\tau^\zeta) \\ &\quad \times \left[\int \prod_l da_l q_b(a_l) \sum_{\boldsymbol{\tau}_1, \cdots, \boldsymbol{\tau}_{K-1}} \delta\left(\boldsymbol{\tau} \prod_l \boldsymbol{\tau}_l; 1 \right) \prod_{l=1}^{K-1} P(\boldsymbol{\tau}_l|a_l) \right]^{C-1} \end{aligned} \tag{5.88}$$

$$\begin{aligned} W_{\rm c}(h|\boldsymbol{\tau}, a) &= \sum_{\rho} p_{\rm chan}\left(\rho|\tau^1\right) \\ &\quad \times \int \prod_{r=1}^{C-1} du_r \prod_{r=1}^{C-1} Z_{\rm c}\left(u_r|\boldsymbol{\tau}\right) \delta\left(h - \Psi_{\rm v}\left(\{u_r\}_{r=1..C-1} | g(\rho) + f(a) \right) \right), \\ Z_{\rm c}\left(u|\boldsymbol{\tau}\right) &= \frac{\int \prod_l da_l q_b(a_l) \sum_{\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \cdots, \boldsymbol{\tau}_{K-1}} \delta\left(\boldsymbol{\tau} \prod_l \boldsymbol{\tau}_l; 1\right) \prod_l P(\boldsymbol{\tau}_l|a_l)}{\int \prod_l da_l q_b(a_l) \left(\sum_{\boldsymbol{\tau}'_1, \boldsymbol{\tau}'_2, \cdots, \boldsymbol{\tau}'_{K-1}} \delta\left(\boldsymbol{\tau}' \prod_l \boldsymbol{\tau}'_l; 1\right) \prod_l P(\boldsymbol{\tau}'_l|a_l) \right)} \\ &\quad \times \int \prod_{l=1}^{K-1} dh_l W_{\rm c}(h_l) \delta\left(u - \Psi_{\rm c}\left(\{h\}_{l=1..K-1} \right) \right), \end{aligned} \tag{5.89}$$

with $\Psi_{\rm c}$ and Ψ_v given by the BP algorithm (5.37) and (5.38). In order to take the limit $g \to 0$ in (5.80) we make the following replica symmetric assumptions on the

$\boldsymbol{\tau}$-dependencies

$$P(\boldsymbol{\tau}|a) = \int dm\ X_{\rm c}(m|a)\ \prod_{\zeta=1}^{g} \mathcal{Q}\left(\tau^{\zeta}|m\right), \tag{5.90}$$

$$\begin{aligned} W_{\rm c}(h|\boldsymbol{\tau},a) &= \int dz\ P(z|\boldsymbol{\tau},a)\ W_{\rm c}(x|\tau^1,z,a) \\ &= \frac{1}{P(\boldsymbol{\tau}|a)} \int dz\ \theta(z|a)\ P(\boldsymbol{\tau}|z,a)\ W_{\rm c}(h|\tau^1,z,a) \\ &= \left[\int dm\ X_{\rm c}(m|a)\ \prod_{\zeta=1}^{g} \mathcal{Q}\left(\tau^{\zeta}|m\right)\right]^{-1} \\ &\quad \times \int dz\ \theta(z|a)\ W_{\rm c}(h|\tau^1,z,a) \prod_{\zeta=1}^{g} \mathcal{Q}\left(\tau^{\zeta}|z\right), \end{aligned} \tag{5.91}$$

with $\int dz\,\theta(z|a) = 1$ and $\int dh\, W_{\rm c}(h|\sigma,a,z) = 1$. The distribution $X_{\rm c}(m|a)$ fulfills the self-consistent equation

$$X_{\rm c}(m|a) = \int \prod_{r=1}^{C-1} dv_r Y_{\rm c}\left(v_r\right) \delta\left(m - \Psi_{\rm v}\left(\{v_r\}_{r=1..C-1}|f(a)\right)\right), \tag{5.92}$$

$$Y_{\rm c}(v) = \int \prod_{l=1}^{K-1} da_l\ q_b(a_l) \int \prod_{l=1}^{K-1} dm_l X_{\rm c}(m_l|a_l) \delta\left(v - \Psi_{\rm c}\left(\{m\}_{l=1..K-1}\right)\right), \tag{5.93}$$

The distribution $\theta(z|a)$ turns out to be also a solution of equations (5.92-5.93). The distributions $W_{\rm c}(h|\tau,z,a)$ are given through the equations

$$W_{\rm c}(h|\tau,m,a)X_{\rm c}(m|a) = \int \prod_r dv_r Y_{\rm c}(v_r) \int \prod_r du_r Z_{\rm c}(u_r|\tau,v_r) \sum_\rho p_{\rm chan}(\rho|\tau)$$
$$\times\delta\left(m - \Psi_{\rm v}\left(\{v_r\}_{r=1..C-1}|f(a)\right)\right)\delta\left(h - \Psi_{\rm c}\left(\{v\}_{l=1..K-1}\right)\right), \qquad (5.94)$$

$$Z_{\rm c}(u|\tau,v)Y_{\rm c}(v) = \int \prod_l da_l\, q_b(a_l) \int \prod_l dm_l X_{\rm c}\left(m_l|a_l\right)$$
$$\times \sum_{\tau_1,\cdots,\tau_{K-1}} P\left(\{\tau_l\}_{l=1..K-1}|\tau,\{m_l\}_{l=1..K-1}\right) \int \prod_l W_{\rm c}\left(h_l|\tau_l,m_l,a_l\right)$$
$$\times\delta\left(u - \Psi_{\rm c}\left(\{h\}_{l=1..K-1}\right)\right)\delta\left(v - \Psi_{\rm c}\left(\{m\}_{l=1..K-1}\right)\right), \qquad (5.95)$$

with

$$P(\{\tau_l\}|\tau,\{z_l\}) \;=\; \frac{\delta\left(\tau\prod_l\tau_l;1\right)\prod_l \mathcal{Q}(\tau_l|z_l)}{\sum_{\tau_1,\cdots,\tau_{K-1}}\delta\left(\tau\prod_l\tau_l;1\right)\prod_l\mathcal{Q}(\tau_l|z_l)}. \qquad (5.96)$$

Equations (5.93) and (5.94) are the density evolution equations for a biased source connected with an asymmetric channel. The distributions $W_{\rm c}(h|\tau,m,a)$ and $Z_{\rm c}(u|\tau,v)$ denote respectively the density of cavity fields on the edges of the Tanner graph given by the BP-algorithm (5.37) and (5.38). The order parameter (5.36) becomes

$$\rho = \int da q_b(a) \int dm X(m|a) \sum_\tau p_{\rm s}\left(\tau|a\right) \int dh W(h|\tau,m,a)\mathrm{sign(h)}, \qquad (5.97)$$

with W and X the real distributions found by substitutions $C-1 \to C$ and $K-1 \to K$. The density evolution equations simplify considerably when we consider unbiased sources $b = 0.5$: we have the stable solution $X_{\rm c}(m|a) = \delta(m)$,

$Y_{\rm c}(v) = \delta(v)$, $W_{\rm c}(h|\tau, z, a) = W_{\rm c}(h|\tau)$ and $Z_{\rm c}(u|\tau, v) = Z_{\rm c}(u|\tau)$ with

$$\begin{aligned} W_{\rm c}(h|\tau) &= \sum_{\rho} p_{\rm chan}\left(\rho|\tau\right) \\ &\quad \times \prod_{r=1}^{C-1} \int du_r Z_{\rm c}\left(u_r|\tau\right) \delta\left(h - \Psi_v\left(\left\{h_r\right\}_{r=1..C-1}|g(\rho)\right)\right), \quad (5.98) \\ Z_{\rm c}(u|\tau) &= \sum_{\tau_1,\cdots,\tau_{K-1}} \frac{\delta\left(\tau_1 \cdots \tau_{K-1}\tau; 1\right)}{2^{K-2}} \\ &\quad \times \int \prod_{\ell=1}^{K-1} dh_\ell W_{\rm c}(h_\ell|\tau_\ell)\, \delta\left[u - \Psi_{\rm c}\left(\left\{h_\ell\right\}_{\ell=1..K-1}\right)\right]. \quad (5.99) \end{aligned}$$

The equations (5.98-5.99) are the density evolution equations found in [172]. We take the limit $\kappa \to 0$ by rescaling the fields $h \to \lambda h$ and $u \to \lambda u$, with $\lambda = -\frac{1}{4}\log(\kappa)$. We decompose the fields through $h = \lambda(h_0 + \lambda^{-1}h_1$ and $u = \lambda(u_0 + \lambda^{-1}u_1$. The distributions of (h_0, u_0) are given by

$$\begin{aligned} W_0(h_0|\tau) &= \int \left(\prod_{r=1}^{C-1} du_r Z_0(u_r|\tau)\right) \sum_{\rho} P_{\rm chan}(\rho|\tau) \\ &\quad \times \delta\left(h - \delta(\rho; 1) - \sum_{r=1}^{C-1} u_r\right), \quad (5.100) \\ Z_0(u|\tau) &= \sum_{\tau_1,\cdots,\tau_{K-1}} \frac{\delta\left(\tau_1 \cdots \tau_{K-1}\tau; 1\right)}{2^{K-2}} \int \left(\prod_{\ell=1}^{K-1} dh_\ell W_0(h_\ell|\tau_\ell)\right) \\ &\quad \times \delta\left(u - \prod_{\ell=1}^{K-1} \mathrm{sign}(\mathrm{h}_\ell)\min\left\{|\mathrm{h}_1|, |\mathrm{h}_2|, \cdots, |\mathrm{h}_{\mathrm{K}-1}|\right\}\right). \quad (5.101) \end{aligned}$$

The equations (5.100) and (5.101) are the DE equations of a BEC with erasure noise $\epsilon = (1+q)/2$. This decoding process corresponds to:

- if $\rho_i = 1$, set $\sigma_i = 1$
- if $\rho_i = -1$ we have no information on the value of σ_i

The joint distributions of (h_0, h_1) and (u_0, u_1) are given by

$$\begin{aligned} W_{\rm c}(h^{(0)}, h^{(1)}|\tau) &= \int \left(\prod_{r=1}^{C-1} du_r^{(0)} du_r^{(1)} Z_{\rm c}(u_r^{(0)}, u_r^{(1)}|\tau) \right) \sum_{\rho} P_{\rm chan}(\rho|\tau) \\ &\quad \times \delta \left(h^{(0)} - \delta(\rho;1) - \sum_{r=1}^{C-1} u_r^{(0)} \right) \\ &\quad \times \delta \Bigg(h^{(1)} - \delta(\rho;-1) g(\rho) \\ &\qquad -\delta(\rho;1) \sum_{\sigma} \sigma\tau \log \left((1-q)^{(\sigma+1)/2} q^{(1-\sigma)/2} \right) - \sum_{r=1}^{C-1} u_r^{(1)} \Bigg), \end{aligned} \tag{5.102}$$

$$\begin{aligned} Z_{\rm c}(u^{(0)}, u^{(1)}|\tau) &= \sum_{\tau_1, \cdots, \tau_{K-1}} \frac{\delta(\tau_1 \cdots \tau_{K-1}\tau; 1)}{2^{K-2}} \int \left(\prod_{\ell=1}^{K-1} dh_\ell W_{\rm c}(h_\ell^{(0)}, h_1^{(0)}|\tau_\ell) \right) \\ &\quad \times \delta \left[u^{(0)} - \left(\prod_{\ell=1}^{K-1} \mathrm{sign}(\mathrm{h}_\ell^{(0)}) \right) \min \left\{ |\mathrm{h}_1^{(0)}|, |\mathrm{h}_2^{(0)}|, \cdots, |\mathrm{h}_{\mathrm{K}-1}^{(0)}| \right\} \right] \\ &\quad \times \delta \left[u^{(1)} - \prod_{\ell=1; \ell \notin \mathcal{K}}^{K-1} \mathrm{sign} \left(\mathrm{h}_\ell^{(0)} \right) \mathrm{atanh} \left(\prod_{\ell \in \mathcal{K}} \tanh \left(\mathrm{h}_\ell^{(1)} \right) \right) \right], \end{aligned} \tag{5.103}$$

with $\mathcal{K}$ the set of indices with $h_\ell = \min \left\{ |\mathrm{h}_1^{(0)}|, |\mathrm{h}_2^{(0)}|, \cdots, |\mathrm{h}_{\mathrm{K}-1}^{(0)}| \right\}$.

The free energy

Substitution of the ansätze (5.86), (5.87), (5.90) and (5.91) in the expression (5.81) of the free energy leads to, after taking the limits $n \to 0$ and $g \to 0$,

$$\begin{aligned} -f_{\rm RS} = &\left(\frac{C}{K}(K-1) \right) \mathbb{E}_b^{(K)} \left[\Delta F_{\rm RS}^{(K)} (\{h_l\}) \right] \\ &-\mathbb{E}_b^{(1)} \left[\Delta F_{\rm RS}^{(1)} (\{h_l^r\}; g(\rho) + f(a)) \right], \end{aligned} \tag{5.104}$$

with

$$\mathbb{E}_b^{(K)}\left[g\left(\{h_l\}\right)\right] = \int \prod_{l=1}^{K} da_l dm_l dh_l W_{\rm c}\left(h_l, m_l, a_l\right)$$
$$\times \sum_{\tau_1,\cdots,\tau_K} P\left(\{\tau_l\}\,|\,\{m_l\}\right) g\left(\{h_l\}\right)$$
$$\mathbb{E}_b^{(1)}\left[f(\{h_l^r\}\,; g(\rho)+f(a))\right] = \int da\; q_b(a)$$
$$\times \int \prod_{r=1}^{C}\prod_{l=1}^{K-1} da_l^r dm_l^r dh_l^r W_{\rm c}\left(h_l^r, m_l^r, a_l^r\right)$$
$$\times \sum_{\tau} p_a(\tau) \prod_{r=1}^{C} \sum_{\tau_1^r\cdots\tau_{K-1}^r} P\left(\{\tau_l^r\}\,|\tau, \{z_l^r\}\right)$$
$$\times \sum_{\rho} p_{\rm chan}(\rho|\tau,a) f(\{x_l^r\}\,; g(\rho)+f(a)), \tag{5.105}$$

and where we used the abbreviations

$$P(\{\tau_l\}\,|\,\{z_l\}) = \frac{\delta\left(\prod_l \tau_l; 1\right)\prod_l \mathcal{Q}\left(z_l^r|\tau_l\right)}{\sum_{\tau_1,\cdots,\tau_K}\delta\left(\prod_l \tau_l; 1\right)\prod_l \mathcal{Q}\left(z_l^r|\tau_l\right)}, \tag{5.106}$$

$$\Delta F_{\rm RS}^{(K)} = -\frac{1}{\beta}\log\left(1+\prod_{l=1}^{K}\tanh\beta h_l\right) + \frac{1}{\beta}\log(2), \tag{5.107}$$

$$\Delta F_{\rm RS}^{(1)} = -\frac{1}{\beta}\log\left(\sum_{\tau} e^{\beta(g(\rho)+f(a))\tau}\prod_{r=1}^{C}\frac{1}{2}\left(1+\tau\prod_{l=1}^{K-1}\tanh\beta h_l^r\right)\right), \tag{5.108}$$

and

$$W_{\rm c}(h,m,a|\tau) \equiv W_{\rm c}(h|\tau,m,a)X_{\rm c}(m|a)q_b(a), \tag{5.109}$$

$$Z_{\rm c}(u,v,a) \equiv Z_{\rm c}(u|v,a)Y_{\rm c}(v)q_b(a). \tag{5.110}$$

The free energy reads in the unbiased case of $b = 0.5$:

$$-f_{\rm RS} = \left(\frac{C}{K}(K-1)\right)\mathbb{E}_{\rm RS}^{(K)}\left[\Delta F_{\rm RS}^{(K)}\left(\{h_l\}\right)\right] - \mathbb{E}_{\rm RS}^{(1)}\left[\Delta F_{\rm RS}^{(1)}\left(\{h_l^r\}\,;g(\rho)\right)\right], \tag{5.111}$$

with

$$\mathbb{E}_{\rm RS}^{(K)}\left[g\left(\{x_l\}\right)\right] = \left(\sum_{\tau_1,\cdots,\tau_K}\frac{\delta\left(\prod_l \tau_l;1\right)}{2^{K-1}}\int\prod_{l=1}^{K}dh_l\,W_{\rm c}(h_l|\tau_l)\right)g\left(\{h_l\}\right), \tag{5.112}$$

$$\mathbb{E}_{\rm RS}^{(1)}\left[g(\{x_l^r\}\,;h)\right] = \sum_{\tau}\left(\prod_{r=1}^{C}\sum_{\tau_1^r\cdots\tau_{K-1}^r}\frac{\delta\left(\tau\prod_l\tau_l^r;1\right)}{2^{K-2}}\right)$$
$$\times\int\prod_{r,l}dh_l^r W_{\rm c}(h_l^r|\tau_l^r)\int dh\,p(h|\tau,\frac{1}{2})\,g\left(\{h_l^r\}\,;g(\rho)\right). \tag{5.113}$$

The expression of the free energy in the unbiased case is clearly much simpler. Generalization of these formulas to irregular graphs is straightforward and important for the design of codes [147]. We note that $\Delta F_{\rm RS}^{(K)}$ and $\Delta F_{\rm RS}^{(1)}$ correspond, in the framework of the cavity method [105], respectively, to the free-energy shifts due to link- and site-addition.

5.5.4 Random codeword model

We consider the limit of regular LDPC codes for $K, C \to \infty$ with the rate $R = 1 - \frac{C}{K}$ fixed. This limit, implying for each bit an infinite number of parity checks, is of little practical importance. Nevertheless, it already contains a wealth of information about the code's performance. It will give us a first flavor about the effects of asymmetry in Gallager codes. We consider only the case of unbiased sources $b = 0.5$.

Definitions

When $K, C \to \infty$ the codewords $\{\tau_i\}_{i=1..M} \in \mathcal{C}$, for an unbiased source, are sampled with a uniform probability [114]. In this limit we obtain the *random codeword model* (RCM), equivalent to random coding discussed in subsection 5.2. The RCM contains a set of 2^N codewords $\left\{\tau_i^{(c)}\right\}_{i=1..M}$, with $c = 0..(2^N - 1)$.

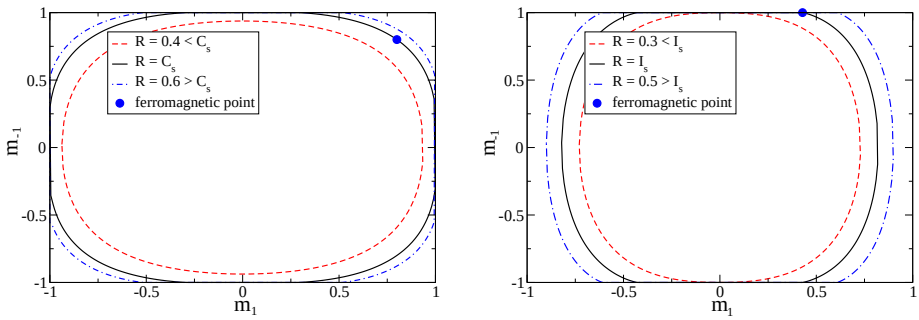


Figure 5.8: The zero entropy lines, $s(m_1, m_{-1}) = 0$, for different rates R at $q = 0$ and $p = 0.4$ in the unbiased case of $b = \frac{1}{2}$. The entropy is positive within the encircled regions. For rates $R > I(\kappa q, q; 1/2)$ the entropy is positive and decoding is not possible. Left: BSC with $C_{\rm s} = I\,(q, q; 1/2)$. Right: Z-channel with $I_{\rm s} = I\,(0, q; 1/2)$.

These codewords are quenched variables drawn from a uniform distribution. The codeword $c = 0$ is the sent codeword. The Hamiltonian of the RCM is

$$H\left(c|\left\{\rho_i\right\}_{i=1..M}\right) = -\sum_{i=1}^{M} g(\rho_i)\tau_i^{(c)}, \tag{5.114}$$

with $c = 1..2^N$ indicating the codeword. The quenched variables ρ_i are dependent on $\tau_i^{(0)}$ through $p_{\rm chan}$ equation (5.64). We call the state $\left\{\tau_i^{(0)}\right\}_{i=1..M}$ the F state.

The Shannon limit

We calculate the number of codewords with a given energy e. If there is only one codeword with $e = e_{\rm F}$, the energy of the F state, we know that errorless decoding by calculation of the energy. We write down

$$H(\gamma) = -\sum_{i=1} |g(\rho_i)|\tau_i^{(0)}\tilde{\tau}_i^{(\gamma)}, \tag{5.115}$$

with $\tilde{\tau}_i^{(\gamma)} = \tau_i^{(\gamma)}\mathrm{sign}\left(\mathrm{g}(\rho_{\rm i})\right)\tau_{\rm i}^{(0)}$. We forget about the tilde in $\tau_i^{(\gamma)}$ in the sequel. We have made two transformations:

- We make a gauge transformation from codewords to the overlap of the codewords with the original codeword sent

$$\tau_i^{(\gamma)} \to \tau_i^{(0)} \tau_i^{(\gamma)}, \tag{5.116}$$

such that the ferromagnetic codeword corresponds with

$$\left\{\tau_i^{(0)}\right\}_{i=1..M} = (1, 1, \ldots, 1). \tag{5.117}$$

- We make the transformation

$$\tau_i^{(\gamma)} \to \text{sign}\left(g(\rho_i)\right) \tau_i^{(\gamma)}. \tag{5.118}$$

After these transformations the energy of a codeword $\left\{\tau_i^{(\gamma)}\right\}$ becomes

$$\frac{H(\gamma)}{N} = \sum_{l=\pm 1} \epsilon_l(\gamma) = -\sum_{l=\pm 1} |H_l| m_l(\gamma), \tag{5.119}$$

with $m_l(\gamma) = \frac{1}{N_l} \sum_{i=1}^{M} \delta(|h_i|, |H_l|) \sigma_i^{(\gamma)}$, $N_l = \sum_{i=1}^{M} \delta(|h_i|, |H_l|)$ and $H_l = g(l)$. Since the codewords are random, the entropy for a given m_1 and m_{-1} is given by

$$s\left(m_1, m_{-1}\right) = (R-1)\log 2 + \left(\frac{1-q+p}{2}\right) Q(m_1) + \left(\frac{1+q-p}{2}\right) Q(m_{-1}), \tag{5.120}$$

with $Q(m) = -\sum_{\lambda=\pm 1} \frac{1}{2}(1+\lambda m) \log[\frac{1}{2}(1+\lambda m)]$. The threshold for MAP decoding is given by the noise levels (p^*, q^*) where $s(m_1^F, m_{-1}^F) = 0$, with (m_1^F, m_{-1}^F) the magnetizations of the F state:

$$m_1^F = \text{sign}(H_1) \frac{1-q-p}{1-q+p}, \qquad m_{-1}^F = \text{sign}(H_{-1}) \frac{1-q-p}{1+q-p}. \tag{5.121}$$

This zero entropy condition corresponds to $R = I(p^*, q^*; 1/2)$, with I the mutual information for an asymmetric channel, see equation (5.65). As expected, we find back the Shannon limit, see figure 5.8.

The phase diagram

In finite temperature decoding we restrict the energies ϵ_1 and ϵ_{-1} by introducing the Lagrange parameters β_1 and β_{-1}. The free energy $f(\beta_1, \beta_{-1})$ is defined through the Legendre transformation

$$-\left(\frac{\beta_1+\beta_{-1}}{2}\right) f(\beta_1,\beta_{-1}) \quad = \quad \max_{\epsilon_1,\epsilon_{-1}} \left(s(\epsilon_1,\epsilon_{-1}) - \beta_1\epsilon_1 - \beta_{-1}\epsilon_{-1}\right).$$

The system contains three states:

- *the P state*: This state corresponds with the "sea" of random codewords. The state has a positive entropy s. It is the solution $W(h|\tau) = \delta(h)$ to the DE equations (5.98-5.99) for $C \to \infty$. Since this solution is always stable no decoding through the BP algorithm possible. The paramagnetic free energy f_P is given by

$$\begin{aligned} -\beta f_P(\beta_1,\beta_{-1}) \quad &= \quad R\log(2) + \left(\frac{1-q+p}{2}\right)\log\cosh\beta_1 H_1 \\ &\quad + \left(\frac{1-p+q}{2}\right)\log\cosh\beta_{-1}H_{-1}. \end{aligned} \tag{5.122}$$

- *the SG state*: This states has a zero entropy. It consists of a subexponential amount of codewords in the system size M. The condensation of the system into the SG state is signalized by the entropy. The entropy becomes negative since the partition sum is dominated by atypical states. The number of these states becomes zero when $M \to \infty$. This corresponds with an entropy crisis as found in the REM, see subsection 2.4.4. Analogously to the P-SG phase transition in the REM the system condensates on the codewords having the ground state energy. This can be checked through a calculation of the participation ratio equation 2.151. The SG state has a free energy $f_{\rm SG}$ given by

$$f_{SG}(\beta_1,\beta_{-1}) \quad = \quad f_P(\beta_1^f,\beta_{-1}^f), \tag{5.123}$$

  with $s(\beta_1^f,\beta_{-1}^f) = 0$.

- *the F state*: this state contains one codeword, i.e. $\left\{\tau_i^{(0)}\right\}_{i=1..M}$. In this state thermodynamic decoding is successful. The free energy of the F state is

$$-\beta f_{\rm F} \quad = \quad \beta H_1\left(\frac{1-q-p}{1-q+p}\right) + \beta H_{-1}\left(\frac{1-q-p}{1+q-p}\right). \tag{5.124}$$

The P-SG transition takes place at a temperature $(\beta_1, \beta_{-1}) = (\beta_1^f, \beta_{-1}^f)$, with

$$\frac{1-q^*+p^*}{2} H\left[\frac{1-q^*-p^*}{1-q^*+p^*}\right] + \frac{1+q^*-p^*}{2} H\left[\frac{1-p^*-q^*}{1+p^*-q^*}\right]$$
$$= \frac{1-q+p}{2} H\left[\frac{(1-q)^{\beta_1^f} - p^{\beta_1^f}}{(1-q)^{\beta_1^f} + p^{\beta_1^f}}\right] + \frac{1+q-p}{2} H\left[\frac{(1-p)^{\beta_{-1}^f} - q^{\beta_{-1}^f}}{(1-p)^{\beta_{-1}^f} + q^{\beta_{-1}^f}}\right]. \tag{5.125}$$

The F-SG phase transition is given by $R = I(p, q; 0.5)$. The triple point lies at $(\beta_1, \beta_{-1}) = (1, 1)$. Comparing the free energies of the F, P and SG state we find for $\beta_1 = \beta_{-1}$ the phase diagram presented in figure 5.9. From the phase diagram of the RCM we conclude that increasing the degree of asymmetry in the channel noise increases the robustness of the system to noise. In asymmetric channels one can decode errorless for much higher noise values.

5.5.5 Dynamic and thermodynamic decoding in regular LDPC codes

There are some crucial differences between the phase diagrams of FiC codes and the phase diagram of the RCM. Decoding through the BP algorithm is possible for FiC codes since the P phase is not always stable, see figure 5.4.

The different phases

LDPC contain four different phases: the ferromagnetic (F) phase, the sub-optimal (S) phase, a spin-glass (SG) phase stable to RSB and a SG phase unstable to RSB. The F phase and the SG phase stable to RSB are similar to the corresponding phases found in the RCM. The S phase is similar to the P phase in the RCM. The SG phase which is unstable to RSB only exists in FiC systems. This has been checked through a high connectivity expansion [114, 129].

- *F-phase*: The DE equations (5.94) and (5.95) admit the solution $W(h|\tau, m, a) = \delta(h - \infty)$ and $Z(u) = \delta(u - \infty)$ corresponding with the codeword sent. The

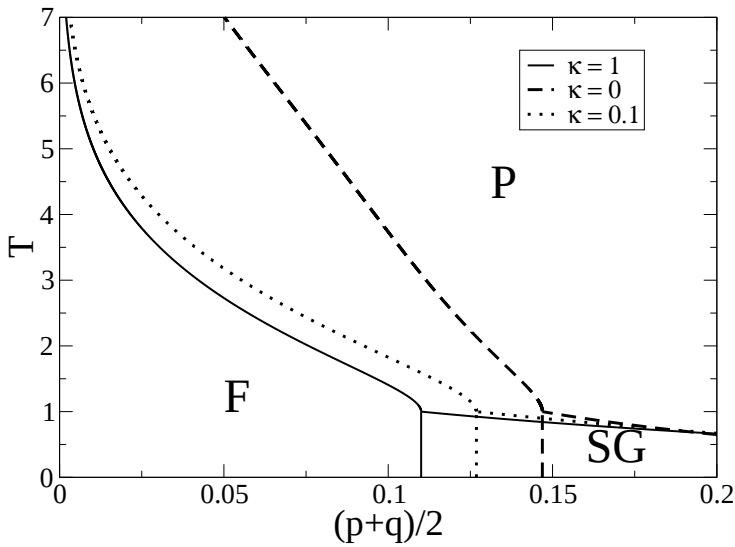


Figure 5.9: The (T,(p+q)/2)-phase diagram for the random codeword model with a rate $R = 1/2$ for different degrees of the channel asymmetry $\kappa = p/q$. Solid lines indicate the thermodynamic transitions to the paramagnetic (P), spin glass (SG) or ferromagnetic (F) phases.

entropy in the F phase is zero. The energy and the free energy equal

$$f_{\rm F} = e_{\rm F} = -\int da\, q_b(a) \int \prod_{r,l} da_l^r\, q_b(a_l^r) \int \prod_{r,l} dm_l^r\, X_{\rm c}(m_l^r|a_l^r)$$

$$\times \sum_{\tau} p_a(\tau) \prod_{r=1}^{C} \sum_{\tau_1^r \cdots \tau_{K-1}^r} P\left(\{\tau_l^r\}\,|\sigma, \{m_l^r\}\right) \sum_{\rho} p_{\rm chan}\left(\rho|\tau\right) g(\rho). \tag{5.126}$$

The F-phase corresponds to errorless decoding and thus reliable communication.

- *S-phase*: In this phase the order parameter (5.36) has a value $0 < \rho < 1$. This phase contains a positive entropy and is thus determined by an exponential amount of codewords. This phase corrupts BP at low values of $\beta < \beta_{\rm SG}$.

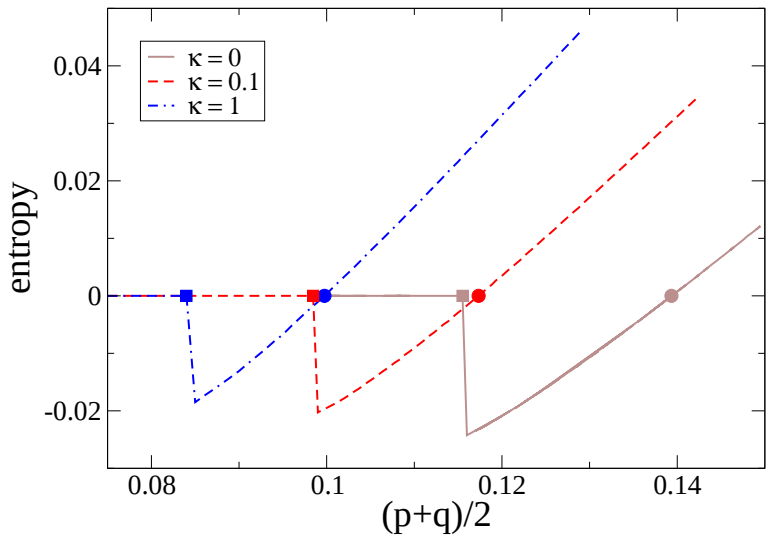


Figure 5.10: The RS entropy s_{RS} of the stationary solution of the order-parameter equations (5.98)-(5.99), as a function of $(p+q)/2$, from the initial state $W(h|\tau) = \delta(h)$ for a BAC at various degrees of asymmetry κ. The lines are calculated at the Nishimori temperature for a regular (3,6) code. The SG solution emerges at a certain noise level when the RS entropy becomes negative (square markers). This solution corrupts the BP algorithm. The F solution dominates the thermodynamic averages. The S state appears when the entropy becomes positive (circle markers). This phase dominates the thermodynamic average.

- *SG-phase stable to local RSB*: In this phase the free energy is dominated by a subexponential amount of codewords. The BP equations converge. Thi2.2.3s phase is responsible for corrupting BP for values $1 \leq \beta < \beta$SG.
- *SG-phase unstable to local RSB*: The free energy is dominated by a subexponential amount of codewords and the BP-equations do not converge. This phase will corrupt BP for values $\beta < 1$

Threshold values

We consider two types of threshold values, *dynamic* and *thermodynamic* ones.

- *The dynamic threshold value*, also called the spinodal threshold value, is defined as the maximal noise value that allows successful decoding with the BP equations. The value $q_{\rm d}(\beta)$ is the largest value of the noise such that the F phase is the only stable solution. Since the decoding algorithm starts in the paramagnetic state $W(h|\tau, m, a) = \delta(h)$ and $Z(u|v, a) = \delta(u)$, the presence of any other stable solution to the BP equations will distort the decoding process. In figure 5.4 the dynamic threshold is shown graphically as the appearance of a S phase which distorts the decoding process.

- *The thermodynamic threshold value* is defined as the maximal noise value such that the order parameter ρ is one. The order parameter ρ is given by

$$\rho = M^{-1}\overline{\sum_{i=1}^{M}\Big(\tau_i \text{sign}\langle\sigma_{\rm i}\rangle_{\beta,\mathbb{H},\{\rho_{\rm j}\}_{\rm j=1..M}}\Big)}, \tag{5.127}$$

  where the average $\langle\cdot\rangle_{\beta,\mathbb{H},\{\rho_j\}_{j=1..M}}$ is taken with respect to the distribution $P_{\mathbb{H},\beta}$ equation (5.47) and the overline is an average over the quenched variables $\{\rho_j\}_{j=1..M}$, $\mathbb{H}$ and $\{\tau_j\}_{j=1..M}$. For $\beta = 1$ the thermodynamic threshold value corresponds with the threshold for MAP decoding. Indeed, as $e(\beta = 1) = e_{\rm F} = f_{\rm F}$ the free energy $f_{\rm S} = f_{\rm F}$ when $s_{\rm S} = 0$. In figure 5.10 we show the entropy at $\beta = 1$ as a function of the noise value for different degrees of the asymmetry κ. We see that the entropy becomes negative at the dynamic threshold value (square markers). The MAP threshold is given by the value of $(q + p)/2$ where the entropy s becomes again zero (circle markers). In figure (5.4) the thermodynamic threshold value is visualized as the point where the free energy value $f_{\rm S} < f_{\rm F}$.

We calculate the threshold values with the population dynamics algorithm, see subsection 2.2.3. In table 5.1 we compare the different threshold values $q_{\rm c}(\beta = 1)$ and $q_{\rm d}(\beta = 1)$ for different degrees of asymmetry in the channel and different regular Gallager codes. In figure 5.11 we present the values of the dynamic and thermodynamic threshold values in the (q, p)-phase and compare them with the Shannon limit for $\beta = 1$ and for $R = 1/2$. The thermodynamic thresholds approach the limit following from the information rate when $C \to \infty$. The dynamic thresholds go to zero when $C \to \infty$. There is a trade-off between dynamic and thermodynamic decoding. For biased sources we show in figure 5.12 the threshold values calculated with a modified version of the population dynamics algorithm subsection 2.2.3 which allows to solve the equations (5.92- 5.95), see [129]. One approaches closer to the Shannon limit with biased sources. The dynamic threshold values are presented in figure 5.13 for unbiased sources as a function of $T(\beta^{-1})$ and $(p+q)/2$. Asymmetry in the channel increases the range of reliable communication. As expected the BP equations perform best when the true noise value is known, i.e. $\beta = 1$. Below $\beta = 1$ the decoding threshold decreases corresponding with

		$\kappa = 1$		$\kappa = 0$		$\kappa = 0.1$	
C	K	q_{d}	q_{c}	q_{d}	q_{c}	q_{d}	q_{c}
5	6	0.13739(5)	0.26436(3)	0.35546(2)	0.70400(5)	0.28709(2)	0.54800(1)
3	4	0.16703(1)	0.20959(1)	0.45580(2)	0.57591(4)	0.35426(1)	0.44167(1)
4	6	0.11692(1)	0.17245(1)	0.30802(2)	0.47132(5)	0.24615(1)	0.36373(2)
3	6	0.08406(1)	0.09972(1)	0.23146(2)	0.27880(1)	0.17977(2)	0.21329(2)
4	8	0.07681(1)	0.10717(1)	0.20056(2)	0.2905(1)	0.16137(2)	0.22635(2)

Table 5.1: The spinodal (q_{d}) and thermodynamic (q_{c}) critical noise levels of regular LDPC codes for the BAC calculated within the RS-ansatz at $T = 1$ and with an unbiased source. The variable $\kappa = p/q$ controls the amount of symmetry in the channel noise. The thresholds for the Z-channel are calculated with $\kappa \sim \mathcal{O}(10^{-8})$.

a reentrance effect. For $C > 2$ and $K > 2$ the dynamic and thermodynamic transitions are discontinuous in the order parameter ρ. For $C = 2$ the transitions are continuous and thus $q_{\mathrm{d}} = q_{\mathrm{c}}$.

5.5.6 The spin-glass phase

In this subsection we elaborate on the spin-glass phase. This phase distorts the dynamic decoding process when T is small. The SG phase contains two important features: condensation and clustering.

Condensation

In general p-spin models, see subsection 2.4.2, a temperature T_{d} a transition occurs from a P phase to a 1RSB-phase and at a temperature $T_{\mathrm{c}} < T_{\mathrm{d}}$ there is an entropy crisis corresponding with the vanishing of the configurational entropy σ. In the present work, because of the infinitely strong interactions, the first phase transition will not take place [98]. There is an entropy crisis as in the RCM and in the REM.

At such an entropy crisis the RS value of the entropy s_{RS} becomes negative. The thermodynamic average $\overline{f}$ is dominated by atypical states that disappear for $M \to \infty$. The typical value of the free energy f is given by

$$f(\beta) = \begin{cases} f_{\mathrm{RS}}(\beta) & \text{when } s_{\mathrm{RS}} \geq 0 \\ f_{\mathrm{RS}}(\beta_f) & \text{when } s_{\mathrm{RS}} < 0 \end{cases}, \tag{5.128}$$

with β_f the inverse temperature at which the system freezes in the lowest energy non-ferromagnetic configuration, i.e. $s_{\mathrm{RS}}(\beta_f) = 0$.

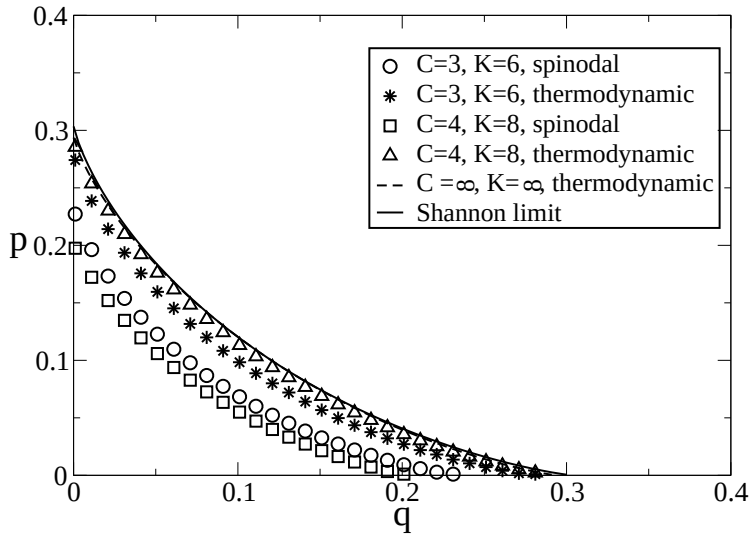


Figure 5.11: Two Gallager codes for BAC's of rate $R = 1/2$ are compared. Critical noise level lines in the (p, q)-parameter space for $\beta = 1$ and an unbiased source are shown. The thresholds obtained through the channel capacity, and for dense codes ($(C, K) = (\infty, \infty)$) through the symmetric information rate, are almost indistinguishable.

Because of the hard constraints, i.e. the delta functionq in equations (5.48) and (5.84), we have indeed that the following frozen ansatz, see [114],

$$P(\boldsymbol{\sigma}|\boldsymbol{\tau}, a) = \sum_{\{n^\gamma\}} \left(\int dh\, W_m(h|\boldsymbol{\tau}, a) \prod_{\gamma=1}^{n/m} \mathcal{Q}(n^\gamma|h) \right) \left(\prod_{\gamma=1}^{n/m} \prod_{\alpha=1}^{m} \delta\left[\sigma_{\gamma,\alpha}; n^\gamma\right] \right), \tag{5.129}$$

with $m \in [0, 1]$, fulfils the selfconsistent equations (5.84) and (5.85). Using this ansatz in the self-consistent equations (5.85) and the free-energy expression (5.81), we find back the RS-equations (5.94) and (5.104) with $\beta \to \beta m$. The extremization condition $\frac{\partial f_{\mathrm{RS}}}{\partial m} = 0$ corresponds to the zero-entropy condition, which for $m \in [0, 1[$ can only be fulfilled when $\beta \geq \beta_f$. When $\beta < \beta_f$ the entropy does not have a zero value for $m \in [0, 1]$. We take $m = 1$, because there the free energy is indeed

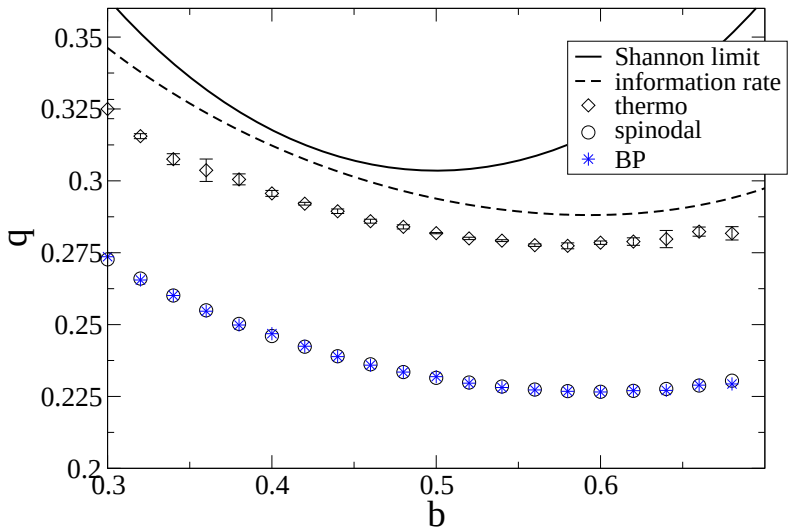


Figure 5.12: Thresholds of the noise level q as a function of the bias for Z-channels with an encoding strategy $(C, K) = (3, 6)$ and $\beta = 1$. The thermodynamic and spinodal lines obtained through population dynamics are compared with thresholds determined through the information rate, the Shannon limit and the results obtained by BP. The dashed and solid lines are the same as the upper dashed and solid lines in figure 5.6.

maximal. This corresponds with the frozen scenario of (5.128). One can recover the results of the frozen ansatz (5.129) within a full 1RSB calculation as done in subsection 2.2.4. One has the solution

$$W_{\rm c}^{\rm 1RSB}[P|\sigma] \quad = \quad \int dw\, W(w|\sigma)\; \delta_F\Big(P(h) - w\delta\left(h-\infty\right) - (1-w)\delta\left(h+\infty\right)\Big), \tag{5.130}$$

to the full 1RSB equations, see [129]. The interpretation is clear: all the codewords are solutions to the cavity equations as a set of infinitely large fields and are therefore the pure states of our system. The variable w denotes the fraction of codewords with a positive spin on a certain site.

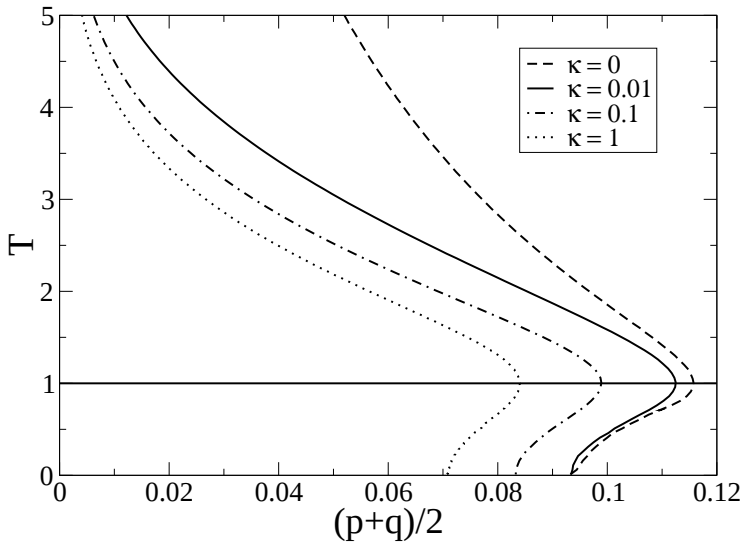


Figure 5.13: The $(T, (p+q)/2)$-phase diagram for the spinodal transition lines within the RS-approximation for a $(C, K) = (3, 6)$ regular Gallager code on a BAC with an unbiased source. The variable $\kappa = p/q$ controls the amount of asymmetry in the channel noise. The Nishimori line, $T = 1$, is shown.

In figure 5.14 we plot the full thermodynamic phase diagram of a regular $(3, 6)$-Gallager code for a BAC with an unbiased source in the space of $(T, \frac{1}{2}(p + q))$-variables and for three different levels of asymmetry in the channel noise. At the Nishimori temperature one obtains a maximal performance as was already observed in subsection 5.5.2. The re-entrance effects below the Nishimori temperature have disappeared because of the frozen ansatz. One observes again a large improvement in reliable communication over asymmetric channels.

Clustering

From a two replica approach along the lines of subsection 2.2.4 we find that the BP equations stop converging at low temperatures. This indicates local symmetry breaking effects at low temperatures. In [129] we have done a full 1RSB study of the order-parameter equations (5.87). We have used $\mathcal{O}(1e+3)$ distributions of

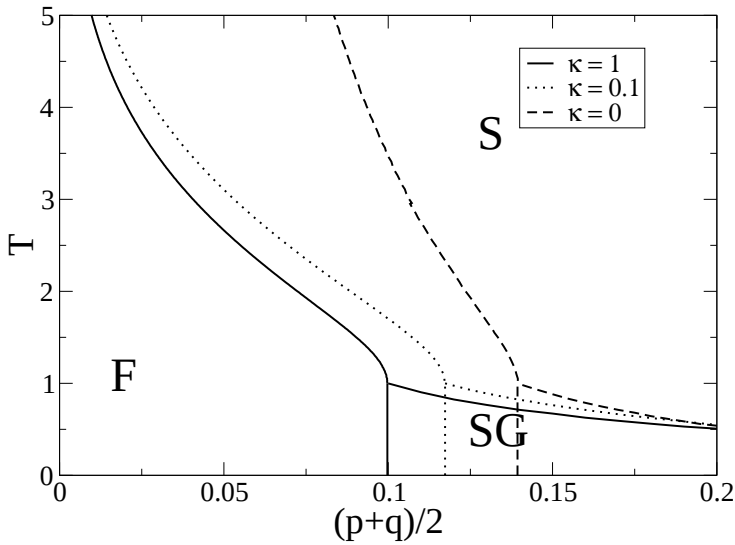


Figure 5.14: The thermodynamic $(T, (p+q)/2)$-phase diagram in the frozen ansatz for a regular $(3, 6)$-regular Gallager code on a BAC for different degrees of the asymmetry κ and an unbiased source. A F, S and SG phase occur.

each $\mathcal{O}(1e+3)$ fields. We interpret the complexity as a function of the entropy for a $(3, 6)$-code on a BSC at $q = 0.09$, $\kappa = 1$ and $T = 1/2$ presented in figure 5.15. The complexity is calculated at the marked point ($\bullet$) in figure 5.17. From the thermodynamical relation between $m = \frac{\mu}{\beta}$ and σ equation (2.110), we know that σ must decrease as a function of m. Eliminating the branches where the complexity increases as a function of μ we find the results in figure 5.15. We find a regime with a positive complexity and a negative entropy. This means that there are an exponential number of solutions to the BP equations and thus the BP algorithm does no longer converge. We also remark that these solution have a negative entropy consistent with the freezing/condensation transition. The complexity is zero at the cross ($\times$) in figure 5.17.

It would be interesting to look for the change of the dynamic thresholds between the replica symmetric and 1RSB algorithms. In order to exclude finite size effects, we would need larger system sizes to determine accurately these thresholds. A valuable addition to our results would be the complexity as a function of

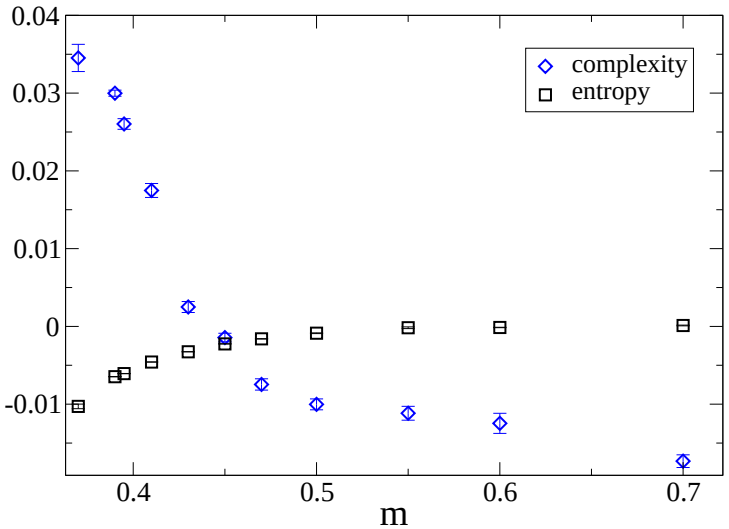


Figure 5.15: The complexity σ (blue markers) and the entropy s (black markers) as a function of the replica symmetry breaking parameter $m = \mu/\beta$ for a $(3, 6)$-regular Gallager code on a BSC with an unbiased source at $p = 0.09$ and $T = \frac{1}{2}$, which corresponds to the marked point ($\bullet$) in the phase diagram of figure 5.17. Error bars of the different data points are shown.

the participation ration (2.152) since this order parameter quantifies the low-temperature phase.

5.5.7 Full phase diagrams

We conclude this section with a full overview of all the phases in LDPC error-correcting codes in figures 5.16-5.17 for, respectively, the Z channel and the BSC channel. The solid lines denote the thermodynamic phase transitions between the F, S and SG state. In the F state the thermodynamic value of the order parameter ρ (5.36) is one and error-less decoding through thermodynamic averages is possible. The dotted lines mark the dynamic phase transitions. The F-SG and F-S transition are determined by points where the SG and S solutions appear and distort the decoding process through the BP algorithm. The transitions between the S and

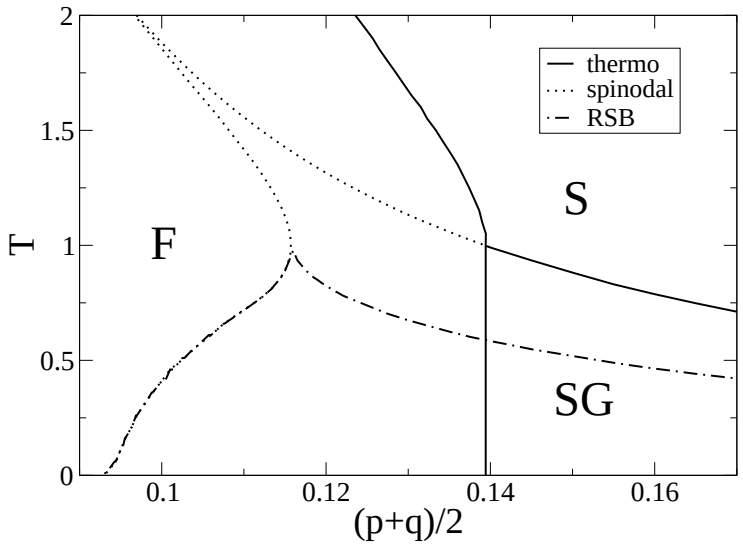


Figure 5.16: Full $(T, (p+q)/2)$-phase diagram for a (3,6)-encoding scheme over a Z-channel. The solid lines indicate thermodynamic phase transitions and the dotted lines represent spinodal transitions. The dashed-dotted line bounds the region of the SG phase which is unstable to RSB. In this region the BP equations do not converge.

SG state are determined by the condensation of the S state in a subexponential number of codewords. The entropy becomes zero at the transition line between the S and SG state. At last the figures show the region where the BP equations do not converge. This region is enclosed by the dotted-striped line and is located at low temperatures. For these parameters there are exponential many solutions to the BP equations. At the Nishimori temperature $\beta = 1$ there are no RSB effects. The BP equations converge always to one solution at the Nishimori temperature. The RSB line reaches the Nishimori temperature at $q_{\mathrm{d}}(\beta = 1)$.

5.6 LDPC on channels with memory

A common problem in modern mobile telecommunication systems is that the strength of the signal varies over time as a result of e.g. the motion of the receiver

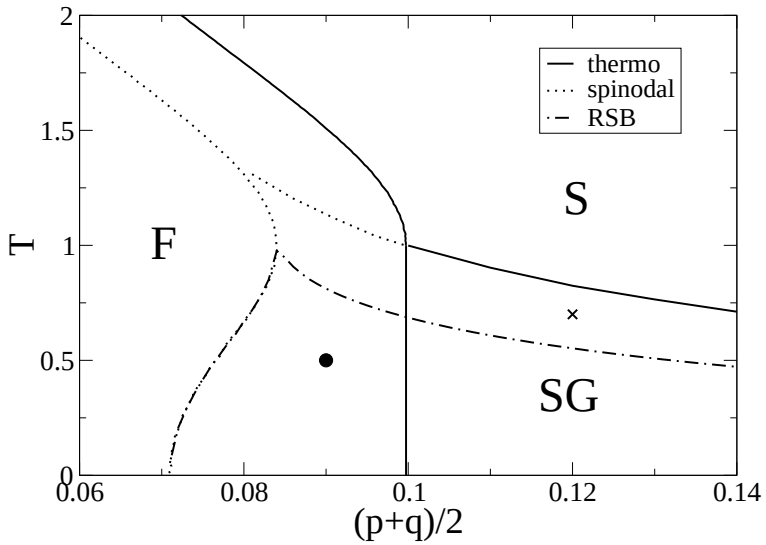


Figure 5.17: The same lines as presented in figure 5.16 for a BSC and a (3,6)-encoding scheme. Two additional points are added where a full 1RSB-calculation is performed. At the cross we have found no 1RSB solution with a positive complexity. At the marked bullet we have found the solution presented in figure 5.15.

with respect to the source and the varying number of obstacles that shadow the signal over time. Channels describing communication of attenuated signals are termed 'fading channels'. Fading channels are modeled by finite-state Markov channels (FSMC) [173]. These channels have fueled significant research activity (for a recent review on the subject see [152]). In FSMCs there exist a number of different channel states that correspond to the various possible attenuation factors. Each of the states describes a memoryless channel characterized by an error probability, while the transition from one state to another occurs according to a stationary Markov process. Since there are different states in the FMSC the error-probabilities between subsequent uses of the channel are correlated, i.e. there is memory in the channel.

Since channels with memory have a higher capacity [125, 67] one would like to introduce memory in the decoding process. Important therefore are the extensions of turbo codes and LDPC codes to FSMCs [62, 44]. The derivation of the density

evolution equations to determine the decoding thresholds [147, 146] is a first step in the design of capacity reaching codes over channels with memory.

Our work is based on techniques that were developed to analyze macroscopic properties of 'small-world' networks. These systems, due to their close relation with real-world networks, have been the subject of intense study from a variety of scientific disciplines [174, 41, 4]. Small-world lattices have a particular architecture that allows both a high clustering coefficient and a small shortest path-length (unlike the random Erdös-Renyi graphs). They are constructed by superimposing random and sparse graphs with a finite average connectivity onto a one-dimensional ring. An exact analysis of the thermodynamic properties of such systems can be found in [132]. As it is, FSMCs can be mapped to small-world lattices, whereby messages between parity checks and codeword-nodes propagate along the sparse graph while messages between channel-state nodes propagate along the one-dimensional chain.

In this section, we present a general method to derive the density evolution equations for symmetric or asymmetric FSMCs. This includes an exact analysis of the Gilbert-Elliot channel (GEC) [65, 48]. We compute the decoding thresholds for the different channels. For symmetric FSMCs we compare the results to [44] while for memoryless channels to the DE equations (5.98-5.99). Apart from the work of [8], channels with memory, or any other FSMC models, have to our knowledge never been analyzed within statistical physics.

5.6.1 Definitions

The memory in the channel noise is modelled through the dependence of the output of the channel $\{\rho_i\}_{i=1..M}$ on the input through the state variable $\{s_i\}_{i=1..M} \in \mathcal{S}^M$:

$$P\left(\{\rho_i\}_{i=1..M} \,|\, \{s_i\}_{i=1..M}, \{\tau_i\}_{i=1..M}\right) = \prod_{i=1}^{M} \Big(P_{\text{chan}}(\rho_i|\sigma_i, s_i)\Big) P_{\text{state}}\left(\{s_i\}_{i=1..M}\right). \tag{5.131}$$

The probability of the states $P_{\text{state}}\left(\{s_i\}_{i=1..M}\right)$ is governed by a Markov process

$$P_{\text{state}}\left(\{s_i\}_{i=1..M}\right) = P_{\text{state}}(s_1) \prod_{i=1}^{M} \mathcal{W}(s_{i+1}|s_i). \tag{5.132}$$

We will denote by $\{\tau_i\}_{i=1..M}, \{t_i\}_{i=1..M}$, respectively, the *true* codeword and *true* channel state vectors that were realized during the signal communication.

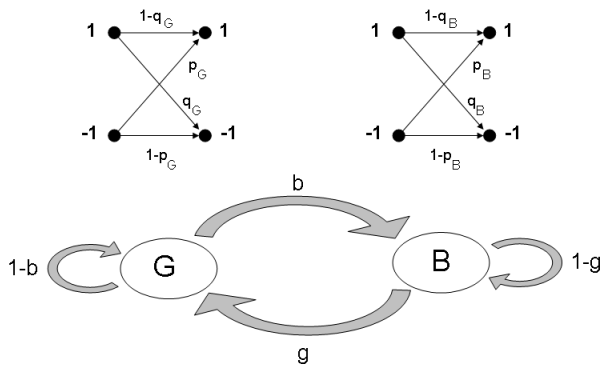


Figure 5.18: A graphical representation of the 2-state FSMC where in each state the channel is a binary asymmetric channel. The bad state B has a higher noise level than the good state G.

Depending on the definition of the Markov process and the channel noise one has different FSMCs. The derivation of the DE equations stays mainly the same. We consider two-state Markov-modulated binary channels. For these channels the noise is a random variable drawn from the distribution

$$P_{\text{chan}}(\rho_i = -\sigma_i | s_i, \sigma_i) = \begin{cases} q_B, & s_i = B,\ \sigma_i = 1 \\ p_B, & s_i = B,\ \sigma_i = -1 \\ q_G, & s_i = G,\ \sigma_i = 1 \\ p_G, & s_i = G,\ \sigma_i = -1 \end{cases} . \tag{5.133}$$

The channel has two states: $\mathcal{S} = \{G, B\}$. Since we take $(p_B + q_B) > (p_G + q_G)$, B is called the bad state and G is called the good state. The Markov process is determined by the transition probability $\mathcal{W}$ given by

$$\mathcal{W} \equiv \begin{bmatrix} 1-b & b \\ g & 1-g \end{bmatrix}, \tag{5.134}$$

with g the transition probability from state B to G and b the transition probability from state G to B (figure 5.18). We define the memory μ_ℓ at time step ℓ of the Markov process as

$$\mu_\ell \equiv \mathcal{W}\left[s_\ell = s | s_0 = s\right] - \mathcal{W}\left[s_\ell = s | s_0 \neq s\right]. \tag{5.135}$$

From (5.134) we find $\mu_\ell = (1-g-b)^\ell \equiv \mu^\ell$ with the time index $\ell = 1, 2, \ldots$ and $\mu \in [-1, 1]$. For $\mu > 0$ we have persistent memory: the probability of remaining in a given state is higher than the steady-state probability of being in that state. For $\mu < 0$ we have an oscillatory memory. We also define the *good-to-bad ratio* $\rho = \frac{g}{b}$. The FSMCs we consider are determined by the 6-tuple $\mathcal{T} = (\mu, \rho, p_B, q_B, p_G, q_G)$. The GEC [48, 65] corresponds to the subset of channels $\mathcal{T}_{\mathrm{GEC}} = (\mu, \rho, p_B, p_B, p_G, p_G)$. We will also consider channels $\mathcal{T}_{\mathrm{AS}} = (\mu, \rho, \kappa q_B, q_B, \kappa q_G, q_G)$ with $\kappa \in [0, 1]$ and $\mathcal{T}_{\mathrm{Z}} = (\mu, \rho, 0, q, q, 0)$.

5.6.2 Density evolution equations

The starting point for the derivation of the DE equations is the calculation of the generating function f of the posterior probability distribution of the codeword $\{\sigma_i\}_{i=1..M}$ given the channel's output $\{\rho_i\}_{i=1..M}$ and the parity-check matrix $\mathbb{H}$:

$$f\left(\{\rho_i\}_{i=1..M}, \mathbb{H}\right) \equiv -\lim_{M\to\infty} M^{-1} \log \sum_{\{\sigma_i\}_{i=1..M}} P\left(\{\sigma_i\}_{i=1..M} \,|\, \{\rho_i\}_{i=1..M}, \mathbb{H}\right).$$

Using Bayes' law and (5.131) we obtain

$$\begin{aligned} &P\left(\{\sigma_i\}_{i=1..M} \,|\, \{\rho_i\}_{i=1..M}, \mathbb{H}\right) \\ &\quad = \frac{P\left(\{\rho_i\}_{i=1..M} \,|\, \{\sigma_i\}_{i=1..M}\right) P\left(\{\sigma_i\}_{i=1..M} \,|\mathbb{H}\right)}{P\left(\{\rho_i\}_{i=1..M} \,|\mathbb{H}\right)} \\ &\quad = Z^{-1} p_{\mathrm{init}}(\{\sigma_i\}_{i=1..M}) \delta_{\mathbb{H}}\left[\{\sigma_i\}_{i=1..M}\right] \\ &\quad \times \sum_{\{s_i\}_{i=1..M}} P_{\mathrm{state}}(\{s_i\}_{i=1..M}) P_{\mathrm{chan}}(\{\rho_i\}_{i=1..M} \,|\, \{s_i\}_{i=1..M}, \{\sigma_i\}_{i=1..M}), \end{aligned}$$

with $P\left(\{\sigma_i\}_{i=1..M} \,|\mathbb{H}\right)$ the initial probability distribution of the codewords and Z a normalisation constant. We will consider unbiased sources of i.i.d.r.v:

$$p_{\mathrm{init}}\left(\{\sigma_i\}_{i=1..M}\right) = 2^{-M}. \tag{5.136}$$

$P_{\mathrm{chan}}\left(\{\rho_i\}_{i=1..M} \,|\, \{s_i\}_{i=1..M}, \{\sigma_i\}_{i=1..M}\right)$ gives the *a priori* probability distribution of the output $\{\rho_i\}_{i=1..M}$ given the state vector $\{s_i\}_{i=1..M}$ and the codeword $\{\sigma_i\}_{i=1..M}$ (5.133). The Kronecker delta restricts the summation only to those codewords that obey the parity-check equation.

Averaging the generating function over the ensemble of parity-check matrices, true-states, true-codewords and outputs gives

$$
\begin{aligned}
-\overline{f} = \lim_{M\to\infty} \frac{1}{M} \sum_{\mathbb{H},\{t_i,\tau_i,\rho_i\}_{i=1..M}} & P(\mathbb{H}) P_{\text{state}}\left(\{t_i\}_{i=1..M}\right) \\
& \times P_{\text{chan}}\left(\{\rho_i\}_{i=1..M} \,|\, \{t_i,\tau_i\}_{i=1..M}\right) \delta_{\mathbb{H}}\left[\{\tau_i\}_{i=1..M}\right] \\
& \times \log\left(\sum_{\{\sigma_i,s_i\}_{i=1..M}} P_{\text{state}}\left(\{s_i\}_{i=1..M}\right) \right. \\
& \left. \times \delta_{\mathbb{H}}\left[\{\sigma\}_{i=1..M}\right] P_{\text{chan}}\left(\{\rho_i\}_{i=1..M} \,|\, \{s_i,\sigma_i\}_{i=1..M}\right) \right),
\end{aligned}
\tag{5.137}
$$

plus irrelevant constant terms. We take for the probability distribution of the parity-check matrices $P(\mathbb{H})$ the distribution of (C,K)-regular codes given by (5.48) and (5.50). The free energy $\overline{f}$ can then be calculated using the replica trick. This results, for $M \to \infty$, in a saddle point integral. The free energy at the saddle point is given by

$$
-\overline{f} = \lim_{n\to 0} \frac{1}{n} \text{extr}_{P,\hat{P}} \Psi\left(P(\boldsymbol{\sigma},\tau), \hat{P}(\boldsymbol{\sigma},\tau)\right),
\tag{5.138}
$$

with Ψ the exponent of the saddle point integral. The extremization is taken over the order-parameter functions $P(\boldsymbol{\sigma},\tau)$ and $\hat{P}(\boldsymbol{\sigma},\tau)$. These represent the usual order-parameter functions describing finite connectivity systems, see equation (2.24), with $\boldsymbol{\sigma} = (\sigma^1,\ldots,\sigma^n) \in \{-1,1\}^n$ originating from the replication of the dynamic codeword-variables and $\tau \in \{-1,1\}$ stemming from the inclusion of the quenched true codeword in the order function.

The exponent Ψ reaches a minimum at the values $\left(P(\boldsymbol{\sigma},\sigma), \hat{P}(\boldsymbol{\sigma},\sigma)\right)$ that satisfy the saddle point equations:

$$
\begin{aligned}
\hat{P}(\boldsymbol{\sigma},\tau) \;=\; & \sum_{(\boldsymbol{\sigma}_1,\tau_1),\cdots,(\boldsymbol{\sigma}_{K-1},\tau_{K-1})} \prod_{r=1}^{K-1} P(\boldsymbol{\sigma}_r,\tau_r) \\
& \times \delta\left(\tau_1\tau_2\cdots\tau_{K-1}\tau, 1\right) \prod_{\alpha=1}^{n} \delta\left(\sigma_1^\alpha \sigma_2^\alpha \cdots \sigma_{K-1}^\alpha \sigma^\alpha, 1\right),
\end{aligned}
\tag{5.139}
$$

$$
P(\boldsymbol{\sigma},\tau) \;=\; \frac{\text{Tr}\left[V^{N-1}\left(\hat{P}\right) Q\left(\boldsymbol{\sigma},\tau;\hat{P}\right)\right]}{\text{Tr}\left[V^N\left(\hat{P}\right)\right]},
\tag{5.140}
$$

where Q is a matrix with elements:

$$\langle \boldsymbol{s}, t|Q(\boldsymbol{\sigma},\tau;\hat{P})|\boldsymbol{s}',t'\rangle = \left(\hat{P}(\boldsymbol{\sigma},\tau)\right)^{C-1} \mathcal{W}\left[t'|t\right] \prod_{\alpha} \mathcal{W}\left[\left(s'\right)^{\alpha}|s^{\alpha}\right] \prod_{\alpha} P_{\text{chan}}\left(\rho|s^{\alpha},\sigma^{\alpha}\right)\rangle_{\rho|t,\tau}, \quad (5.141)$$

where we used the "Bra-ket" notation. We have introduced the average $\langle\cdot\rangle_{\rho|t,\tau}$ over $P_{\text{chan}}(\rho|t,\tau)$. Note that while the summations over the replicated codeword variables $\{\boldsymbol{\sigma}_i\}_{i=1\ldots M}$ have been performed by reducing the graph into a single-site problem, the summations over the replicated channel-state variables $\{\boldsymbol{s}_i\}_{i=1\ldots M}$ are written as a trace over matrix products in (5.140). This constitutes the key difficulty in our problem as we are dealing with the $(2^n+1)\times(2^n+1)$ replicated transfer matrix:

$$\langle \boldsymbol{s}, t|V(\hat{P})|\boldsymbol{s}',t'\rangle = \sum_{\boldsymbol{\sigma},\tau} \mathcal{W}\left[t'|t\right] \prod_{\alpha} \mathcal{W}\left[\left(s'\right)^{\alpha}|s^{\alpha}\right] \times \left(\hat{P}(\boldsymbol{\sigma},\tau)\right)^{C} \langle \prod_{\alpha} P_{\text{chan}}(\rho|s^{\alpha},\sigma^{\alpha})\rangle_{\rho|t,\tau}. \quad (5.142)$$

To proceed further we have to make an assumption with regard to the structure of the replica space. The simplest, a replica symmetric ansatz, assumes that

$$P(\boldsymbol{\sigma},\tau) = 2^{-\frac{1}{K}} \int dh W_{\text{c}}(h|\tau) \prod_{\alpha} \frac{e^{h\sigma^{\alpha}}}{2\cosh\left(h\right)}, \quad (5.143)$$

$$\hat{P}(\boldsymbol{\sigma},\tau) = 2^{-\frac{K-1}{K}} \int du Z_{\text{c}}\left(u|\tau\right) \prod_{\alpha} \frac{e^{u\sigma^{\alpha}}}{2\cosh\left(u\right)}, \quad (5.144)$$

for some densities W, Z. For the left- and right-eigenvectors $L(\boldsymbol{s},t)$, $R(\boldsymbol{s},t)$ of V we now assume

$$\langle \boldsymbol{s}', t'|R\rangle = \mathcal{P}_R(t') \int dx \Phi_R(x|t') e^{x\sum_{\alpha}\left(s'\right)^{\alpha}}, \quad (5.145)$$

$$\langle L|\boldsymbol{s},s\rangle = \mathcal{P}_L(s) \int dy \Phi_L(y|s) e^{y\sum_{\alpha} s^{\alpha}}. \quad (5.146)$$

The form of these equations follows the RS assumption of [132, 131]. It allows us to take the remaining trace in (5.140). All distributions above are normalized

at $n \to 0$. The densities $\mathcal{P}_R$ and $\mathcal{P}_L$ represent respectively the right- and left-eigenvectors of $\mathcal{W}$:

$$\mathcal{P}_R(t) = \sum_{t'} \mathcal{W}\left[t'|t\right] \mathcal{P}_R(t'), \tag{5.147}$$

$$\mathcal{P}_L(t') = \sum_{t} \mathcal{W}\left[t'|t\right] \mathcal{P}_L(t). \tag{5.148}$$

Following similar computations as in [16, 17, 132], we derive in the limit $n \to 0$ the following self-consistent equations in the densities of messages propagating along the Tanner graph [128]:

$$W_{\rm c}(h|\tau) = \int \left(\prod_{r=1}^{C-1} du_r Z_{\rm c}(u_r|\tau) \right) \int d\zeta M(\zeta|\tau)\, \delta \left[h - \zeta - \sum_{r=1}^{C-1} u_r \right] \tag{5.149}$$

$$\begin{aligned} Z_{\rm c}(u|\tau) &= \sum_{\tau_1,\cdots,\tau_{K-1}} \frac{\delta\left(\tau_1 \cdots \tau_{K-1}\tau; 1\right)}{2^{K-2}} \\ &\quad \times \int \prod_{\ell=1}^{K-1} dh_\ell W_{\rm c}(h_\ell|\tau_\ell)\, \delta \left[u - \operatorname{atanh} \prod_{\ell=1}^{K-1} \tanh(h_\ell) \right], \end{aligned} \tag{5.150}$$

The densities of messages propagating between the Tanner graph and the channel are given by

$$F(\xi|\tau) = \int \left(\prod_{r=1}^{C} du_r Z_{\rm c}(u_r|\tau) \right) \delta \left[\xi - \sum_{r=1}^{C} u_r \right], \tag{5.151}$$

$$\begin{aligned} M(\zeta|\tau) &= 2 \sum_{t,t'} \mathcal{P}_R\left(t'\right) \mathcal{W}\left[t'|t\right] \mathcal{P}_L(t) \\ &\quad \times \int dx dy \Phi_L(y|t) \Phi_R(x|t') \sum_{\rho} P_{\rm chan}(\rho|\tau, t) \\ &\quad \times \delta \left[\zeta - \frac{1}{2} \sum_{\tau} \tau\sigma \log \left(\sum_{ss'} e^{\left(s'x+sy\right)} \mathcal{W}\left[s'|s\right] P_{\rm chan}\left(\rho|\tau, s\right) \right) \right] \end{aligned} \tag{5.152}$$

At last, we have the densities of messages propagating along the channel:

$$\begin{aligned}\Phi_R(x|t) &= \sum_{t'} \mathcal{W}[t'|t] \sum_{\tau} \frac{1}{2} \int d\xi F(\xi|\tau) \int dx' \Phi_R(x'|t') \sum_{\rho} P_{\text{chan}}(\rho|\tau,t) \\ &\quad \times \delta \Bigg[x - \frac{1}{2} \sum_{s} s \log \left(\sum_{s'} \mathcal{W}[s'|s] \frac{e^{x's'}}{2\cosh(x')} \right) \\ &\quad - \frac{1}{2} \sum_{s} s \log \left(\sum_{\sigma} \prod_{r} \frac{e^{\xi\sigma\tau}}{2\cosh\xi} P_{\text{chan}}(\rho|\sigma,s) \right) \Bigg], \end{aligned} \tag{5.153}$$

$$\begin{aligned}\Phi_L(x|t') &= \sum_{t} \frac{\mathcal{W}[t'|t]\,\mathcal{P}_L(t)}{\mathcal{P}_L(t')} \\ &\quad \times \sum_{\tau} \frac{1}{2} \int d\xi F(\xi|\tau) \int dx' \Phi_L(x'|t) \sum_{\rho} P_{\text{chan}}(\rho|\tau,t) \\ &\quad \times \delta \Bigg[x - \frac{1}{2} \sum_{s'} s' \log \Bigg(\sum_{s} \mathcal{W}[s'|s] \frac{e^{x's}}{2\cosh(x')} \\ &\qquad \times \left(\sum_{\sigma} \frac{e^{\xi\sigma\tau}}{2\cosh\xi} P_{\text{chan}}(\rho|\sigma,s) \right) \Bigg) \Bigg]. \end{aligned} \tag{5.154}$$

Equations (5.149-5.154) are the DE equations for the binary asymmetric two-state Markov channel. They describe the evolution of the densities of messages propagating along a tripartite graph figure 5.19. The graph consists of a chain of channel-state nodes connected to codeword nodes and these in turn to parity-check ones. This graphical representation of the decoding process corresponds to an efficient algorithm [43], equivalent to the sum-product algorithm used in channels without memory, see subsection 5.3.3. The tripartite graph has three different sets of vertices: the set V_{code} of codeword nodes, the set V_{pc} of parity-check nodes and the set V_{chan} of channel-state nodes, see figure 5.19. Due to the presence of memory there are 6 types of messages propagating according to:

Message	From	To
$h_{i\to a}$	$i \in V_{\text{code}}$	$a \in V_{\text{pc}}$
$u_{a\to i}$	$a \in V_{\text{pc}}$	$i \in V_{\text{code}}$
$\zeta_{c\to i}$	$c \in V_{\text{chan}}$	$i \in V_{\text{code}}$
$\xi_{i\to c}$	$i \in V_{\text{code}}$	$c \in V_{\text{chan}}$
$x_{R;c+1\to c}$	$c \in V_{\text{chan}}$	
$x_{L;c-1\to c}$	$c \in V_{\text{chan}}$	

The update equations for single-graph instances for these messages (the so-called 'message-passing' equations) [43] correspond to the functions given by the arguments of the delta functions in the DE equations (5.149-5.154). When we compare the equations (5.149-5.154) with the memoryless equations (5.98-5.99) for $p_B = q_B$ and $p_G = q_G$ we notice the following differences:

- The DE equations (5.149) and (5.150) are the same.
- The distribution $M(\zeta|\tau)$ becomes

$$M(\zeta|\tau) \quad = \quad \delta\left[\zeta - \frac{1}{2}\sum_{\tau}\tau\sigma\log\left(P_{\text{chan}}\left(\rho|\right)\right)\right]. \tag{5.155}$$

  The fields ζ are not coupled anymore trough the "chain" of states. The chain of states gets decoupled from the Tanner graph.

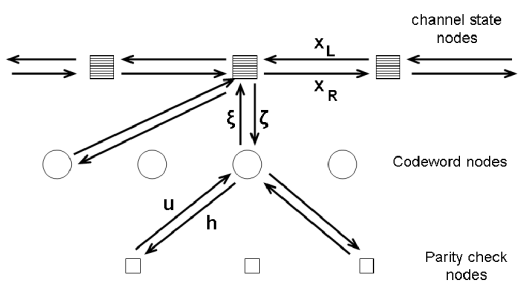


Figure 5.19: The tripartite graph and the messages propagating along the graph for a LDPC code on channels with memory.

5.6.3 Results

We are interested in deriving the critical noise levels beyond which decoding is not possible. This information can be obtained through the observable $\rho_\sigma \equiv$

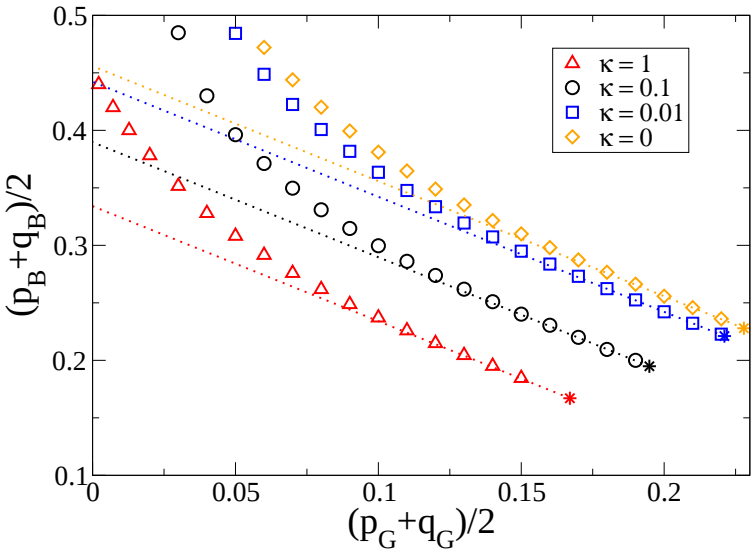


Figure 5.20: Decoding thresholds (markers) of a $(3, 4)$-regular LDPC code on a $\mathcal{T}_{\text{AS}}$-channel presented in the space of $\left(\frac{1}{2}(p_B + q_B), \frac{1}{2}(p_G + q_G)\right)$ for different values of the asymmetry $\kappa \in [0, 1]$. For all symbols the memory equals $\mu = 0.90$ while the good-to-bad ratio is $\rho = 1$. The dotted line represents the memoryless threshold $\mu = 0$. The stars at the end of the markes denote the critical value for memoryless channels where $p_B = p_G$ and $q_B = q_G$.

$\frac{1}{|I_\sigma|} \sum_{i \in I_\sigma} \sigma_i = \int dh W(h|\sigma)\text{sign}(h)$ where I_σ describes the sublattice $I_\sigma = \{i \in V_{\text{code}} | \sigma_i^0 = \sigma\}$ and $\tilde{W}$ the distribution of the marginals of the decoding variables

$$W(h|\sigma) = \int \left(\prod_{r=1}^{C} du_r Z_c(u_r|\sigma) \right) \times \int d\zeta M(\zeta|\sigma)\, \delta \left[h - \zeta - \sum_{r=1}^{C} u_r \right] \tag{5.156}$$

The value $\rho_\sigma = 1$ corresponds to perfect decoding (F phase) while $\rho_\sigma < 1$ describes decoding failure (P phase). We detect the transition by numerically solving the DE equations (e.g. through population dynamics [104]).

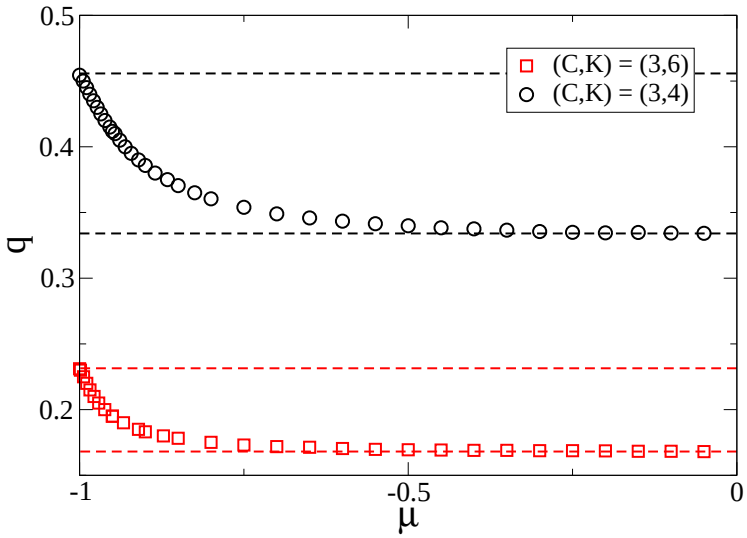


Figure 5.21: Decoding thresholds in the (μ, q) space of (C, K)-regular LDPC codes on a $\mathcal{T}_{\rm Z}$-channel for a good-to-bad ratio $\rho = 1$. The upper dashed lines represent the decoding thresholds for the corresponding memoryless binary asymmetric channel. The lower one corresponds to a value that is twice that of the decoding threshold of a memoryless binary symmetric channel.

The decoding thresholds in the parameter space $\left(\frac{1}{2}(p_G + q_G), \frac{1}{2}(p_B + q_B)\right)$ for a Gallager $(C, K) = (3, 4)$ code on a $\mathcal{T}_{\rm AS}$-channel are shown in figure 5.20. Dotted lines separate ferro- from paramagnetic solutions for memoryless channels with $\mu = 0$, while symbols correspond to channels with memory for $\mu = 0.9$. We show four degrees of channel asymmetry characterized by the variable $\kappa = p_B/q_B = p_G/q_G$. Note that to simplify the presentation of our results the two channel states have here the same κ. The decoding thresholds for $\kappa = 0$ are computed from the DE equations for $\kappa \to 0$, which can be derived when rescaling the fields $h \to \beta h$, $u \to \beta u$, $\zeta \to \beta\zeta$ and $\xi \to \beta\xi$ with $\beta = -\frac{1}{4}\log(\kappa)$. The points marked by the star-symbols at the end of the markers correspond to the points where the two channel states have the same error probability, $p_B = p_G$ and are taken from table 5.1. Beyond the star-symbol (lower-right part of the figure) the roles of the 'good' versus the 'bad' channel are interchanged. In this figure we also see that implementing the memory and the asymmetry of the channel in the decoding

	$\kappa = 1$	$\kappa = 0.1$	$\kappa = 0.01$	$\kappa = 0$
$(p_G + q_G)/2$	$(p_B + q_B)/2$			
0.20	/	0.190(2)	0.242(2)	0.256(2)
0.08	0.262(2)	0.331(2)	0.401(2)	0.420(2)
0.05	0.308(2)	0.396(2)	0.484(2)	/

Table 5.2: Critical noise levels for the (3,4)-regular LDPC code on a $\mathcal{T}_{\text{AS}}$-channel with memory $\mu = 0.9$, good-to-bad ratio $\rho = 1$ and for four degrees of channel asymmetry.

process enhances the threshold values of the decoding process. In the limiting cases of $\kappa = 1$ our results agree very well with those of [44]. In table 5.2 we give the decoding thresholds corresponding to figure 5.20.

In figure 5.21 we present results from the channel $\mathcal{T}_Z$ in which there exist two Z-type states: $(p_B, q_B) = (0, q)$ and $(p_G, q_G) = (q, 0)$ (hence the terms 'good' vs 'bad' are not very meaningful here). This type of configuration can model 'burst-error' channels where a very large number of consecutive bits appear corrupted while the corruption is selective with regards to the input symbol. We show results in the (μ, q) space for Gallager $(C, K) = (3, 4)$ and $(3, 6)$ codes. The lower dashed line corresponds to the noise level $2q_{\text{BSC}}$ where q_{BSC} is the critical level of the memoryless binary-symmetric channel. The fact that the marker at $\mu = 0$ coincides with the dashed line is not a coincidence since in this limit the channel has two complementary Z-type states without memory, and therefore, with an equal transition probability between them. The upper dashed line corresponds to the critical noise level of a memoryless Z-channel [129]. At $\mu = -1$ the transition probabilities become $b = g = 1$ and thus the channel oscillates between the two states. We note that this figure is symmetric with respect to the $\mu = 0$ axis; a property that also follows from the DE equations.

5.7 Conclusion

The density evolution equations determine the performance of LDPC codes when the size of the codewords becomes infinitely large. We have derived the density evolution equations of asymmetric channels and of channels with memory by mapping them on spin models on finitely-connected hypergraphs.

The density evolution equations for asymmetric channels have, in contrast with symmetric channels, a very complicated dependency on the codeword sent, but simplify a lot when the source is unbiased. For unbiased sources we recover the density evolution equations derived through a cavity like approach in [172]. The information rate increases when the asymmetry of the channel

is incorporated in the decoding algorithm. We have studied finite temperature decoding with the belief propagation algorithm and finite temperature decoding through thermodynamic averaging. When the temperature is high the noise is overestimated while for low temperature the noise is underestimated. When the temperature is one the true noise values of the channel is known. We have determined the temperature-noise phase diagram of the model for various degrees of asymmetry and identified the regions where errorless decoding is possible. This region of errorless decoding increases with the degree of asymmetry. At higher noise values the decoding algorithm fails due to a suboptimal phase or a spin-glass phase. The suboptimal phase corresponds with an entropic amount of codewords and appears at high temperatures. The spin glass phase is a condensated phase which appears at low temperatures. The spin-glass phase is very similar to the spin glass phase in the random energy model. At even lower temperatures the message passing algorithm stops converging and clustering phenomena appear.

In the second part of this chapter we have studied LDPC codes for channels with memory. Error-correcting codes on channels with memory are known to outperform the traditional ones on memoryless channels. They can be used in modern mobile communication systems or to model burst-error channels. We have presented a technique for deriving the density evolution equations for multi-state channels. This method originally developed to study 'small-world' systems is based on the diagonalisation of replicated transfer matrices. It turns out that the representation of the LDPC multi-state decoding problem on graphs shares a common architecture with 'small-world' systems: In memoryless channels, decoding occurs with message-passing between symbol variables (the 'spins') which are connected to parity-check variables (the 'couplings'). Channels with memory introduce a new element to this hypergraph which can be seen and treated as a chain of channel-state variables with nearest-neighbor interactions.

We have presented results for the Gilbert-Elliott channel and its generalization to asymmetric two-state channels with memory. The density evolution equations that follow from the analysis reproduce very well the special limiting cases of the Gilbert-Elliott or the memoryless binary-asymmetric channel. The method can be applied to a variety of multi-state error-correcting codes, such as multi-symbol, gaussian-, non-Markovian or intersymbol-interference channels.

5.8 Future prospects: design of codes

Our studies have shown that one can analyze LDPC codes on "complicated" channels involving memory and asymmetry. The design of codes is one of the most interesting aspects of LDPC codes. Optimizing the threshold values $q_{\rm d}(\beta = 1)$ of LDPC codes to the degree distributions v_k and c_l, equations (5.53) and (5.54), one can reach the Shannon limit $q_{\rm d} = q_{\rm s}$. This is proven for erasure channels [93]. For

general symmetric channels one can approach closely the Shannon limit [146] and probably reach it although we are unaware of a proof.

Is it possible to design codes using statistical mechanics methods? To do this one should find a cost function that is optimized at the dynamical transiion $q_{\rm d}$ or behaves monotonously as a function of $q_{\rm d}$. In figure (5.22) we show the dynamical and static thresholds for two sets of irregular codes on a BEC as a function of the degree distribution. Both sets of irregular codes have generating functions $c(x)$ and $v(x)$, equations (5.25) and (5.26), of the form:

$$c(x) = (1-\gamma-\delta)x^3 + \delta x^8 + \gamma x^{30}, \tag{5.157}$$

$$v(x) = \alpha x^3 + (1-\alpha)x^{10}. \tag{5.158}$$

We keep the rate R and the number of links $\overline{k} = 3\alpha + 10(1-\alpha)$ in the Tanner graph fixed. This corresponds, respectively, with a fixed information transfer and a fixed computational efficiency of the decoding algorithm. In the first set of codes $\overline{k} = 5$ while in the second set $\overline{k} = 10$. The rate $R = 0.2$. In figure 5.22 the thresholds $q_{\rm d}$ and $q_{\rm c}$ are plotted as a function of the remaining parameter δ.

The dynamical threshold increases considerably when we consider irregular codes. However, the thermodynamic transition $q_{\rm c}$ is maximal for regular codes. As the thermodynamic transition is linked with minimization of the free energy we do not know which cost function can be used to optimize the dynamical threshold.

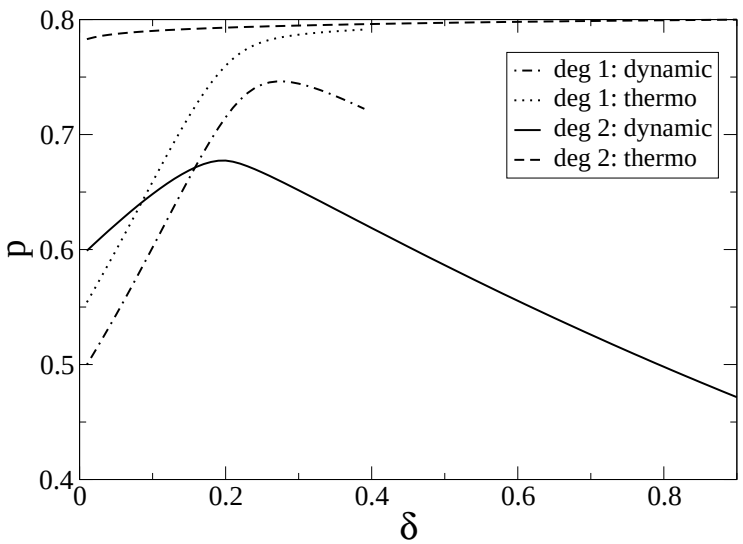


Figure 5.22: The dynamic critical noise level (solid and dashed-dotted lines) and the thermodynamic threshold levels (dashed and dotted lines) for two different irregular Gallager codes on a BEC are shown. The degree distributions are given by the following generating functions: $c(x) = (1 - \gamma - \delta)x^3 + \delta x^8 + \gamma x^{30}$; $v(x) = \alpha x^3 + (1-\alpha)x^{10}$, with a rate $R = 0.2$, $\overline{k} = 5$ for the first code "deg1" and $\overline{k} = 10$ for the second code "deg2". We see that the dynamic threshold is not monotonous in δ and reaches its optimum value for irregular codes. The static transition behaves monotonously in δ and reaches its largest value for regular codes.

Chapter 6

Conclusions

Statistical mechanics is the theory that describes the collective behavior of a large number of interacting molecules, moments, electrons, etc. It is a challenge to extend the underlying concepts and methods to the study of emergent phenomena in complex systems, for instance error-correcting codes, optimization problems on graphs, proteins in a cell or agents trading on the market. A popular approach to address these problems is simplifying them to units interacting via some underlying graph structure. Some examples of these graph structures are metabolic networks or transportation networks. To study processes on these graphs one has to associate a probability distribution on the configuration space of the units, based upon the knowledge one has of the specific underlying biological, transportation or other processes. From an abstract point of view the analysis of these models consists of the study of a probability distribution of spins on a certain graph structure.

When the underling structure is local tree like one can analyze spin models on graphs with the replica method, the cavity method or Monte Carlo methods. Spin models on graphs or finite connectivity systems are studied by message-passing algorithms, such as belief propagation. The belief propagation equations are a central result of the cavity method. These are a set of equations in the marginals of the probability distribution of spins on the given graph structure. From the marginals one can determine relevant thermodynamic quantities such as the energy, the entropy, the magnetization, etc. The density evolution equations are the average over the graph ensemble of the density of messages propagated during the belief-propagation algorithm. They describe the typical behavior of this algorithm.

In this thesis we have studied three open problems within the field of finite connectivity systems:

- **Lévy spin glasses**: The Sherrington-Kirkpatrick model is a fully-connected model for spin glasses. The spin-glass phase consists of a complicated phase space containing infinitely many pure states. The solution of the SK-model has been important in various scientific disciplines where a spin-glass phase is found. The SK-model is universal in the sense that all fully-connected models where the couplings are drawn as i.i.d.r.v. with a finite variance lead to the same phase diagram.

  We have solved a generalization of the SK-model to i.i.d.r.v. with a divergent variance, i.e. a Lévy spin glass. The distribution contains three important parameters: the tail exponent, the skewness and the shift.

  This Lévy spin glass can be solved by decomposing it into a finitely-connected backbone of strong bonds and a background of weak bonds.

  The spin glass phase consists of an exponential number of pure states, similar to the spin glass phase in the SK-model. This is in contrast to the conjectures made in [28]. We find that the SG phase is always unstable with respect to replica symmetry breaking. The ferromagnetic phase contains a mixed state where replica symmetry is unstable. The replica symmetry breaking decreases when the tail exponent increases, i.e. when the system contains more strong bonds and thus less frustration.

- **Parallel dynamics of spin models on graphs**: The low complexity and the accuracy of the belief propagation equations make them a perfect tool to study emergent phenomena of spin models on graphs, given the probability distribution of the spins. In practice one does not always know the distribution of spins. Examples are non-equilibrium models defined through a dynamics which does not fulfill detailed balance. Such problems appear for instance in the activity of neural networks of biological organisms or congestion problems in transportation networks.

  To develop a message-passing algorithm we have considered a simple model: Ising spins evolving through parallel Glauber dynamics on graphs with symmetric and asymmetric links. Evolution of Ising spins on graphs with asymmetric links does not fulfill detailed balance. We have generalized the belief propagation equations to a set of equations in the marginals of the distribution of paths which determines the dynamical evolution of the system. We have taken the stationary limit assuming there are no cycles and neglecting some correlations in time. The limiting cases of fully asymmetric and fully symmetric graphs are exactly solvable. In these cases our algorithm reduces to the known belief propagation equations. For partially asymmetric graphs the algorithm is approximative.

  We have determined the phase diagram for spins interacting through bimodal couplings on a Bethe lattice and a Poissonian lattice. For models containing disorder without frustration, such as fluctuating degrees, we have found a good agreement between the theoretical and simulation results. When

the disorder contains frustration the algorithm needs to be improved. Theoretical results show how the spin glass phase gradually decreases with the degree of asymmetry in the bonds and eventually disappears.

As a second model, we have investigated a neural network for associative memory containing both asymmetry in the couplings and a scale-free degree distribution. These two assumptions make existing models for neural networks more realistic. We find that the correlations between in and out degrees increase the retrieval performance of the neural network considerably.

- **Statistical mechanics of low-density parity-check codes**: Low-density parity-check codes are a family of efficient codes reaching the Shannon limit, i.e. reaching the maximal rate of information to which reliable communication over a noisy channel is possible. The decoding process of low-density parity-check codes can be described through the average over the probability distribution of spins of a relevant order parameter of a finite connectivity system. The mapping of low-density parity-check codes on spin models allows to extend the existing decoding algorithms through the introduction of a temperature. The temperature quantifies the lack of knowledge of the value of the true noise of the channel. A high temperature corresponds with an overestimation of the channel noise while a low temperature corresponds with an underestimation. Low-density parity-check codes on symmetric channels, where the noise is independent of the bits sent, have been thoroughly studied in the literature.

  In this thesis we have made a study of the extension of low-density parity-check codes to asymmetric channels and channels with memory effects in the noise. Asymmetric channels are important in communication through optical fibers while memory effects are important in wireless communication. We show that by taking into consideration the memory and asymmetry of the noise in the decoding process, the decoding thresholds of the corresponding algorithms increase considerably. Therefore, reliable communication at higher rates is possible.

  The statistical mechanics approach leads to two possible decoding algorithms: the first algorithm corresponds with solving the belief-propagation equations on a bipartite Tanner graph and the second algorithm consists of taking a thermodynamic average over the corresponding probability distribution of spins. The temperature-noise phase diagrams of error-correcting codes contain three phases: a ferromagnetic phase corresponding with error free communication, a suboptimal ferromagnetic phase and a spin-glass phase, both corresponding with failure of reliable communication. In the spin-glass phase the system is condensed on a sub-exponential number of erroneous codewords while in the suboptimal phase error correction gets corrupted by an exponential number of erroneous codewords in the codeword length.

When we compare the density evolution equations in the asymmetric case with the density evolution equations of symmetric channels the main difference is the dependence of the densities on the codeword sent.

The belief propagation algorithm on channels with memory is defined on a tripartite graph. Thereby, it is necessary to include a set of messages propagating along the channel. The density evolution equations of this algorithm are found by mapping the corresponding spin model on a small-world hypergraph and by calculating the disorder-averaged free energy of the posterior probability distribution using the method of replicated transfer matrices. Although technically involved, the method is straightforward and can be generalized easily to any channel with memory.

Appendix A

Useful identities

A.1 Gaussian integrals

The multivariate Gaussian integral equals

$$\begin{aligned} I(\mathcal{H}, \boldsymbol{b}) &= \int \left(\prod_{i=1}^{m} dx_i \right) \exp\left(-\frac{1}{2} \sum_{i,j=1}^{m} x_i \mathcal{H}_{ij} x_j + \sum_{i=1}^{m} b_i x_i \right) \\ &= \frac{(2\pi)^{m/2}}{\sqrt{\det \mathcal{H}}} \exp\left[\frac{1}{2} \sum_{i,j=1}^{m} b_i \mathcal{H}_{ij}^{-1} b_j \right], \end{aligned} \tag{A.1}$$

with $\mathcal{H}$ a symmetric matrix with $\mathcal{R}e\,(\lambda_i) \geq 0$ and $\lambda_i \neq 0$, for all eigenvalues λ_i of $\mathcal{H}$.

A.2 The saddle-point method

We want to calculate the following integral:

$$\mathcal{I}_m = \lim_{N \to \infty} \int d\boldsymbol{x} \exp\left(N f(\boldsymbol{x}) \right), \tag{A.2}$$

with $\boldsymbol{x} = (x_1, x_2, \cdots, x_m)$. We consider the Taylor expansion:

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^{(0)}) + \frac{1}{2}\left(\boldsymbol{x} - \boldsymbol{x}^{(0)}\right)^T \mathcal{H}(\boldsymbol{x}^{(0)}) \left(\boldsymbol{x} - \boldsymbol{x}^{(0)}\right) + \mathcal{O}\left(\left(\boldsymbol{x} - \boldsymbol{x}^{(0)}\right)^3\right), \tag{A.3}$$

around a point $\boldsymbol{x}^{(0)}$ with $\frac{\partial}{\partial x_j} f(\boldsymbol{x})\Big|_{\boldsymbol{x}=\boldsymbol{x}^{(0)}} = 0$ for all j. The Hessian $\mathcal{H}(\boldsymbol{x})$ is defined through:

$$\mathcal{H}_{ij}(\boldsymbol{x}) = \frac{\partial^2}{\partial x_j \partial x_i} f(\boldsymbol{x}). \tag{A.4}$$

We make the transformation: $\boldsymbol{x} \to \sqrt{N}^{-1}\boldsymbol{x} + \boldsymbol{x}^{(0)}$ and use the Gaussian integral (A.1) to obtain

$$\mathcal{I}_m = \frac{(2\pi)^{m/2}}{N^{m/2}\sqrt{\det \mathcal{H}}} \exp\left(N f(\boldsymbol{x}^{(0)})\right)\left(1 + \mathcal{O}\left(N^{-1/2}\right)\right), \tag{A.5}$$

if the eigenvalues of $\mathcal{H}$ have a positive real part.

Appendix B

Graphs and Hypergraphs

For an introduction on graph theory we refer to [18, 72]. In this appendix we introduce/summarize all the notations, definitions and graph results used in the thesis.

B.1 Definitions

B.1.1 Undirected graphs

A undirected graph is a couple (V, E) with V the set of vertices and E the set of edges. The number of elements $|V|$ in V is the size of the graph. The elements of E are unordered pairs $(i, j) \in V \times V$ where (i, j) and (j, i) denote the same edge. An edge (i, j) is incident to the two vertices i and j. Two vertices i, j are adjacent when $(i, j) \in E$. We define *simple undirected graphs* through a symmetric traceless $|V| \times |V|$-dimensional connectivity matrix $\boldsymbol{C}$, with $c_{ij} = [\boldsymbol{C}]_{ij} \in \{0, 1\}$, such that $(i, j) \in E \Leftrightarrow c_{ij} = 1$.

A graph can be presented graphically by $|V|$ labelled points. The points with label i and j are connected by a line segment when $c_{ij} = 1$, otherwise they are not connected. The neighbourhood of a vertex ∂_i is the set of all vertices adjacent to i. The degree of a vertex is defined through $k_i \equiv |\partial_i|$. We define the "closed" neighbourhood of a vertex $\overline{\partial}_i \equiv \partial i \cup \{i\}$. In a k-regular graph all vertices have the same degree $k = k_i$. A perfect matching is a graph where every vertex has degree one.

B.1.2 Directed graphs

A directed graph is a couple (V, E) with E now the set of directed edges, i.e. ordered pairs $(i, j) \in V \times V$. We say that two vertices i, j are adjacent when $(i, j) \in E$ or $(j, i) \in E$. We define *simple directed graphs* through a traceless $|V| \times |V|$-dimensional connectivity matrix $\boldsymbol{C}$, with $c_{ij} = [\boldsymbol{C}]_{ij} \in \{0, 1\}$. When $c_{ij} = 1, c_{ji} = 0$ there is a directed edge going from i to j; when $c_{ij} = c_{ji} = 1$ there is a undirected edge between i and j and when $c_{ij} = c_{ji} = 0$ there is no edge between i and j. The set $E^{\mathrm{d}} \subset E$ contains all the directed edges and the set E^{sym} contains all the symmetric edges. Directed graphs are presented through $|V|$ labelled points. A undirected edge is denoted by a line segment while directed edges are presented by arrows. Since the graphs are directed we have four types of neighborhoods and degrees:

- The in-neighbourhood $\partial_i^{\mathrm{in}} \equiv \{j \in V | (j, i) \in E\}$ of the vertex i and its in-degree $k_i^{\mathrm{in}} \equiv |\partial_i^{\mathrm{in}}|$
- The out-neighbourhood $\partial_i^{\mathrm{out}} \equiv \{j \in V | (i, j) \in E\}$ of the vertex i and its out-degree $k_i^{\mathrm{out}} \equiv |\partial_i^{\mathrm{out}}|$
- The symmetric neighbourhood $\partial_i^{\mathrm{sym}} = \partial_i^{\mathrm{in}} \cap \partial_i^{\mathrm{out}}$ of the vertex i and its symmetric degree $k_i^{\mathrm{sym}} \equiv |\partial_i^{\mathrm{sym}}|$
- The neighbourhood $\partial_i = \partial_i^{\mathrm{in}} \cup \partial_i^{\mathrm{out}}$ of the vertex i and its degree $k_i \equiv |\partial_i|$

B.1.3 Hypergraphs

A p-uniform hypergraph is defined as a couple $G = (V, E)$ with $E \subset V^p$, a set of p-tuples. A p-uniform hypergraph is uniquely determined by a connectivity tensor $\boldsymbol{C}^{(p)}$ with elements

$$\left[\boldsymbol{C}^{(p)}\right]_{i_1 i_2 \cdots i_p} = c_{i_1 i_2 \cdots i_p} = \begin{cases} 1 & \text{when } (i_1, i_2, \cdots, i_p) \in E \\ 0 & \text{when } (i_1, i_2, \cdots, i_p) \notin E \end{cases} . \tag{B.1}$$

A non-uniform hypergraph is defined as a couple $G = (V, E)$ with

$$E \subset V \times V \cup V \times V \times V \cup \cdots . \tag{B.2}$$

In this case one can partition $E = E^{(2)} \cup E^{(3)} \cup E^{(4)} \cdots E^{(p_{\max})}$, with $p_{\max}$ the maximum size of the edges $E^{(n)} = E \cap V^n$. A non-uniform hypergraph is determined by $(p_{\max} - 1)$ connectivity tensors $\boldsymbol{C}^{(n)}$. The degree of a vertex is defined through

$$k_i \equiv \sum_j c_{ij} + \sum_{j<k} c_{ijk} + \cdots \tag{B.3}$$

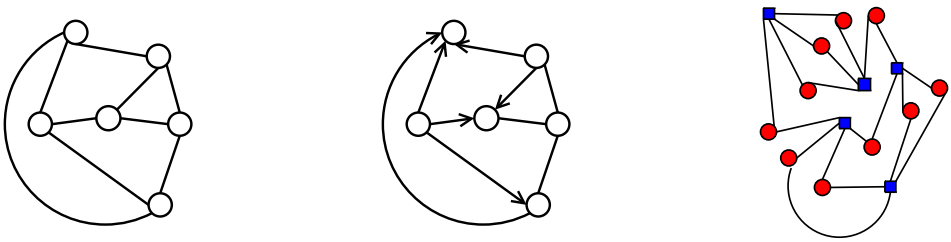


Figure B.1: From left to right we have: a 3-regular simple undirected graph, a simple directed graph and a 4-uniform 2-regular hypergraph.

A hypergraph is k-regular when all vertices have the same degree k.

B.1.4 Bipartite graphs

A bipartite graph is a graph $G = (V, E)$ with $V = V_1 \cup V_2$ and $E \subset V_1 \times V_2$. A bipartite graph is a perfect matching when the degree of every vertex is one.

B.2 Graph ensembles

A graph ensemble is a couple $\mathcal{G} = (P, \mathcal{S})$, with $\mathcal{S}$ a set of graphs and P a probability distribution defined on the set $\mathcal{S}$.

B.2.1 Erdős-Rényi random graphs

The ensemble $\mathcal{G}_{\mathrm{e}}(N, M)$ was studied by Erdős and Rényi in [49]. The set $\mathcal{S}$ consists of all graphs with N vertices and M edges which are present in the ensemble with an equal probability:

$$P_{\mathrm{e}}(G; N, M) \equiv \frac{1}{Z}\delta_K\left(M, \sum_{i<j} c_{ij}\right). \tag{B.4}$$

The normalisation constant Z in (B.4) is given by

$$Z = \begin{pmatrix} N(N-1)/2 \\ M \end{pmatrix}. \tag{B.5}$$

B.2.2 The Poissonian ensemble

The Poissonian ensemble $\mathcal{G}_{\rm p}(N,c)$ consists of graphs drawn with a probability:

$$P_{\mathcal{G}_p}(G;N,c) \equiv \prod_{i<j} \left(\left(1-\frac{c}{N}\right) \delta_K(c_{ij};0) + \frac{c}{N}\delta_K(c_{ij};1) \right). \tag{B.6}$$

For $N \to \infty$ the Poissonian ensemble becomes equivalent to the Erdős-Rényi ensemble. One draws a graph from this ensemble by drawing every edge (i,j) with a probability $p = \frac{c}{N}$.

The two ensembles $P_{\rm e}(G;N,M)$ and $P_{\rm p}(G;N,p)$ are equivalent for $N \to \infty$. Indeed, the probability distribution $P(M)$ is given by

$$P(M) = \exp\left(-\overline{M}\right) \frac{\overline{M}^M}{M!}, \tag{B.7}$$

with $\overline{M} = c(N-1)/2$. The fluctuations $\sigma(M)/\overline{M} = \frac{1}{\sqrt{\overline{M}}}$ for $N \to \infty$. The relative fluctuations of the edges get very small for $N \to \infty$ and $\mathcal{G}_{\rm p}\left(N,p\right) = \mathcal{G}_{\rm e}\left(N, \frac{p}{2}(N-1)\right)$. Therefore, in practice one draws a graph from the ensemble $\mathcal{G}_{\rm p}(N,c)$ through the following process: draw M edges uniformly random from the set $V \times V$.

The degree distribution $p_{\rm d}$ of the Poissonian ensemble equals:

$$p_{\rm d} = \left(\begin{array}{c} N-1 \\ d \end{array} \right) \left[\frac{c}{N}\right]^c \left[1-\frac{c}{N}\right]^{N-d-1}. \tag{B.8}$$

For $N \to \infty$ one has:

$$p_{\rm d}(k) = \frac{c^k \exp\left(-c\right)}{k!}, \tag{B.9}$$

which explains the name of this ensemble. The Poissonian degree has the useful property

$$\sum_{k=0}^{\infty} \frac{p_{\rm d}(k)k}{c} A(k-1) = \sum_{k=0}^{\infty} p_{\rm d}(k) A(k). \tag{B.10}$$

B.2.3 Graphs with a fixed degree sequence

We consider the ensemble $\mathcal{G}_{\rm deg}\left(N, p_{\rm d}(k)\right)$ with N the size of the graph and $p_{\rm d}\left(k\right)$ the asymptotical degree distribution. The probability distribution of these graphs

is defined in two steps:

$$P_{\rm deg}(G|N, p_{\rm d}(k)) = \left(\prod_{i=1}^{N} p_{\rm d}(k_i) \right) P_{\rm deg}(\boldsymbol{C}|\, \{k_i\}_{i=1..N}), \tag{B.11}$$

with

$$P_{\rm deg}(\boldsymbol{C}|\, \{k_i\}_{i=1..N}) \sim \left(\prod_{i<j} \frac{c}{N} \delta(c_{ij}; 1) + \left(1 - \frac{c}{N}\right) \delta\left(c_{ij}; 0\right) \right) \times \prod_{i=1}^{N} \delta\left(k_i; \sum_j c_{ij} \right). \tag{B.12}$$

The constant of proportionality $\mathcal{M} = \sum_{\boldsymbol{C}} P_{\rm deg}(\boldsymbol{C}|\, \{k_i\}_{i=1..N})$, is, for $N \to \infty$, given by:

$$\mathcal{M} = \exp\left[N \left(-\frac{c}{2} + c \log(c) - \sum_{k \geq 0} p_d(k) \log(k!) \right) \right]. \tag{B.13}$$

We draw graphs from this ensemble using the following process:

- Generate N random numbers $\{k_i\}_{i=1..N}$ from $p_{\rm d}(k)$
- Form a set T containing k_i copies of each vertex
- Generate a random matching of the elements in T

A similar ensemble was studied in [112], hence, they are sometimes called *Molloy-Reed random graphs*. For $N \to \infty$ the following conditions are met, see [112]

- The degree sequence of all the graphs is the same up to order $\mathcal{O}(N^{-1})$
- All graphs with a given degree sequence are drawn with the same probability

Some popular ensembles appear repeatedly in the literature:

- The Poissonian ensemble: $p_d(k) = \exp(-c)\, \frac{c^k}{k!}$
- The Bethe ensemble: $p_d(k) = \delta(k; c)$. The graphs drawn from the Bethe ensemble are called Bethe lattices
- Scale-free graphs: in this case the degree distribution is given by a power law $p_d(k) = ak^{-\gamma}$, with a a normalization constant

B.2.4 The directed Poissonian ensemble

The directed Poissonian ensemble $\mathcal{G}_{\rm p}(N,c,\epsilon)$ is defined through the following probability function

$$P_{\rm p}(\boldsymbol{C};c,\epsilon) \equiv \prod_{i<j}\left(\frac{c}{N}\delta(c_{ij};1)+\left(1-\frac{c}{N}\right)\delta(c_{ij};0)\right) \prod_{i>j}\left(\epsilon\delta(c_{ij};c_{ji})+(1-\epsilon)\left(\frac{c}{N}\delta(c_{ij};1)+\left(1-\frac{c}{N}\right)\delta(c_{ij};0)\right)\right). \tag{B.14}$$

The mean connectivity is given by c while ϵ is the fraction of symmetric edges. For $\epsilon = 1$ the ensemble $\mathcal{G}_{\rm p}(N,c,\epsilon)$ reduces to the ensemble $\mathcal{G}_{\rm p}(N,c)$. In the recursive equations (4.23), (4.24) and (4.27) for dynamics we need to know the degree distribution. The degree distribution $p_{\rm p}$, defined through

$$p_{\rm p}(k^{\rm in},k^{\rm out},k^{\rm sym}) \equiv \frac{\sum_i \sum_{\boldsymbol{C}} P(\boldsymbol{C})\delta(k^{\rm in};\sum_j c_{ij})\delta(k^{\rm out};\sum_j c_{ji})\delta\left(k^{\rm sym};\sum_j c_{ij}c_{ji}\right)}{N}, \tag{B.15}$$

is equal to [126]

$$p_{\rm p}(k^{\rm in},k^{\rm out},k^{\rm sym}) = \left(\exp\left[-c\epsilon\right]\frac{(c\epsilon)^{k^{\rm sym}}}{k^{\rm sym}!}\right)\left(\exp\left[-(1-\epsilon)c\right]\frac{(c(1-\epsilon))^{k^{\rm in}-k^{\rm sym}}}{(k^{\rm in}-k^{\rm sym})!}\right) \left(\exp\left[-(1-\epsilon)c\right]\frac{(c(1-\epsilon))^{k^{\rm out}-k^{\rm sym}}}{(k^{\rm out}-k^{\rm sym})!}\right). \tag{B.16}$$

When we sum out the outdegrees we get

$$p_{\rm p}(k^{\rm in},k^{\rm sym}) = \frac{e^{-c}c^{k^{\rm in}}}{k^{\rm in}!}\left(\begin{array}{c} k^{\rm in} \\ k^{\rm sym}\end{array}\right)\epsilon^{k^{\rm sym}}(1-\epsilon)^{k^{\rm in}-k^{\rm sym}}. \tag{B.17}$$

The degree distribution (B.17) has the useful property

$$\sum_{k^{\rm in},k^{\rm sym}} p_{\rm p}(k^{\rm in},k^{\rm sym})\frac{k^{\rm sym}}{c_{\rm sym}}A\left(k^{\rm sym},k^{\rm in}\right) = \sum_{k^{\rm in},k^{\rm sym}} p_{\rm p}(k^{\rm in},k^{\rm sym})A\left(k^{\rm sym}+1,k^{\rm in}+1\right). \tag{B.18}$$

B.2.5 Directed ensemble with a fixed degree sequence

We consider the ensemble $\mathcal{G}_{\text{deg}}\left(N, p_{\text{d}}(k^{\text{in}}, k^{\text{out}}, k^{\text{sym}})\right)$ with N the size of the graph and $p_{\text{deg}}\left(k^{\text{in}}, k^{\text{out}}, k^{\text{sym}}\right)$ the asymptotical joint degree distribution. The probability distribution of these graphs is defined in two steps:

$$\begin{aligned} P_{\text{deg}}\left(G|N, p_{\text{d}}\left(k^{\text{in}}, k^{\text{out}}, k^{\text{sym}}\right)\right) &= \left(\prod_{i=1}^{N} p_{\text{d}}\left(k_i^{\text{in}}, k_i^{\text{out}}, k_i^{\text{sym}}\right)\right) \\ &\times \quad P_{\text{deg}}\left(\boldsymbol{C}|\left\{k_i^{\text{in}}, k_i^{\text{out}}, k_i^{\text{sym}}\right\}_{i=1..N}\right) \end{aligned} \tag{B.19}$$

with

$$\begin{aligned} & P_{\text{deg}}\left(\boldsymbol{C}|\left\{k_i^{\text{in}}, k_i^{\text{out}}, k_i^{\text{sym}}\right\}_{i=1..N}\right) \\ & \sim \prod_{i<j}\left(\frac{c}{N}\delta(c_{ij};1) + \left(1-\frac{c}{N}\right)\delta(c_{ij};0)\right) \\ & \times \prod_{i>j}\left(\epsilon\delta(c_{ij};c_{ji}) + (1-\epsilon)\left(\frac{c}{N}\delta(c_{ij};1) + \left(1-\frac{c}{N}\right)\delta(c_{ij};0)\right)\right) \\ & \times \prod_{i=1}^{N}\delta\left(k_i^{\text{in}};\sum_j c_{ji}\right)\prod_{i=1}^{N}\delta\left(k_i^{\text{out}};\sum_j c_{ij}\right)\prod_{i=1}^{N}\delta\left(k_i^{\text{sym}};\sum_j c_{ij}c_{ij}\right) \end{aligned} \tag{B.20}$$

with $c = \lim_{N\to\infty} N^{-1}\sum_{i,j} c_{ij}$ and $\epsilon = \frac{\sum_{i,j} c_{ij}c_{ji}}{cN}$. It is possible to generate a graph from the ensemble $\mathcal{G}_{\text{deg}}$ with the following algorithm:

1. Draw for every vertex $i \in V$ a triple $(k_i^{\text{in}}, k_i^{\text{out}}, k_i^{\text{sym}})$ from the distribution p_{d}
2. Create k^{sym} sockets on each graph. Draw a random perfect matching on the graph of sockets.
3. Create a bipartite graph which is a perfect matching containing k_i^{in} sockets for each vertex on one side and k_i^{out} sockets for each vertex on the other side.

B.2.6 Local tree structure

It is Important that the graphs have a local tree structure. It is rather easy to show that for $N \to \infty$ there are a finite number $\mathcal{O}(N^0)$ of loops of finite length

$\mathcal{O}(N^0)$. Indeed, for a Poissonian graph the number of k-cycles is:

$$N(N-1)...(N-k+1)\left[\frac{c}{N}\right]^k = c^k + \mathcal{O}(N^{-1}), \tag{B.21}$$

which is $\mathcal{O}(1)$. There are many loops of order $\mathcal{O}(\log(N))$. For a rigorous proof we refer to [18]. But, it is possible to make an easy heuristic argument. Because the graph is looplike we have c^ℓ vertices on a distance ℓ from a root node. Because a graph has N vertices this number is bounded from above by N.

B.2.7 Percolation

Random graphs contain a percolation transition. Below the percolation transition the largest component is of size $\mathcal{O}(\log(N))$ while above the transition there is one giant component with $\mathcal{O}(N)$ vertices. As shown by Erdős-Rényi this transition takes place at $M = \frac{1}{2}N$ [50]. For graphs with a given degree sequence this transition takes place at $\sum_k k(k-2)p_{\mathrm{d}}(k) = 0$.

Appendix C

Stable distributions

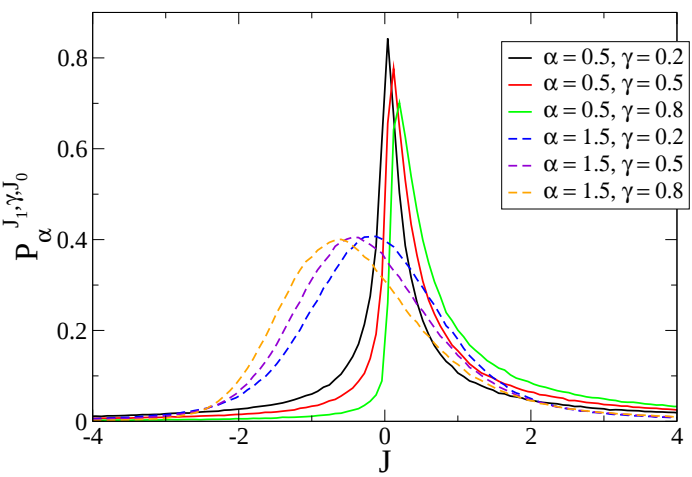


Figure C.1: The Lévy distributions $P_\alpha^{J_1,\gamma,J_0}(J)$ for $J_1 = 1$, $J_0 = 0$ and different values of α and γ. The distributions with $\alpha = 1.5$ approach a Gaussian while the ones for $\alpha = 0.5$ have larger tails. For $\gamma > 0$, the center of the distribution goes to $+\infty$ or $-\infty$ for $\alpha \uparrow 1$ or $\alpha \downarrow 1$, respectively. The coupling distribution fulfills $P_\alpha^{J_1,\gamma,J_0}(J) = P_\alpha^{J_1,-\gamma,J_0}(-J)$.

The purpose of this appendix is to give some intuition on the role of the parameters

α and γ present in stable distributions, defined through eqs. (3.7) and (3.8). Both α and γ are responsible for the shape of the distribution. The main role of the exponent α is to control the decay of the tails. Figure C.1 shows that, for a fixed γ, a decrease in α gives rise to a distribution $P_\alpha^{J_1,\gamma,J_0}(J)$ with larger tails and more sharply peaked around its most probable value J. The center of $P_\alpha^{J_1,\gamma,J_0}(J)$ is also shifted from the negative to the positive x-axis as α decreases from $\alpha > 1$ to $\alpha < 1$. A change of α has no effect on the position of the center when $\gamma = 0$.

The skewness parameter γ controls the relative weight between the positive and negative tails. For $\gamma > 0$, the positive tail of $P_\alpha^{J_1,\gamma,J_0}(J)$ is larger than the negative one; for $\gamma < 0$ vica versa. The distribution is symmetric around J_0 when $\gamma = 0$. For increasing positive values of γ, see figure C.1, the center of $P_\alpha^{J_1,\gamma,J_0}(J)$ shifts to the right or left provided $\alpha < 1$ or $\alpha > 1$, respectively.

C.1 Solution of integrals

The following two expressions are true for any function $f(J)$ for which the integrals on the right hand side exist. A proof of these expressions can be found in the paper [127]. The quantities C_α and Φ are defined in the subsection 3.2.

Expression I

$$\int_{-\infty}^{\infty} \frac{dJd\hat{J}}{2\pi} |\hat{J}|^\alpha e^{i\hat{J}J} f(J) = -\left(\frac{\sqrt{2}}{J_1}\right)^\alpha C_\alpha \int_{-\infty}^{\infty} \frac{dJ}{|J|^{\alpha+1}} \Big[f(J) - f(0)\Big], \quad \text{(C.1)}$$

for $\alpha \in (0,1) \cup (1,2]$.

Expression II

$$\int_{-\infty}^{\infty} \frac{dJd\hat{J}}{2\pi} |\hat{J}|^\alpha \text{sign}(\hat{J}) e^{i\hat{J}J} f(J)$$

$$= \begin{cases} i\left(\frac{\sqrt{2}}{J_1}\right)^\alpha \frac{C_\alpha}{\Phi} \int_{-\infty}^{\infty} \frac{dJ}{|J|^{\alpha+1}} \text{sign}(J) f(J) & \text{if} \quad 0 < \alpha < 1 \\ i\left(\frac{\sqrt{2}}{J_1}\right)^\alpha \frac{C_\alpha}{\Phi} \int_{-\infty}^{\infty} \frac{dJ}{|J|^{\alpha+1}} \text{sign}(J) \Big[f(J) - f'(0)J\Big] & \text{if} \quad 1 < \alpha \leq 2 \end{cases} . \quad \text{(C.2)}$$

Bibliography

[1] The gene ontology consortium, http://www.geneontology.org/.

[2] The mips comprehensive yeast genome database (cygd), http://mips.gsf.de/proj/yeast/cygd/db/.

[3] Digital video broadcasting (dvb); second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications (dvb-s2). Technical Report EN 302 307, ETSI, August 2009.

[4] R Albert and A L Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47, 2002.

[5] R Albert, H Jeong, and A L Barabási. Internet: Diameter of the world-wide web. *Nature*, 401:130–131, 1999.

[6] D J Aldous and A Bandyopadhyay. A survey of max-type recursive distributional equations. *Annals of Applied Probability*, 15:1047, 2005.

[7] D J Amit, H Gutfreund, and H Sompolinsky. Storing infinite numbers of patterns in a spin-glass model of neural networks. *Phys. Rev. Lett.*, 55:1530, 1985.

[8] J A Anguita, M Chertkov, M A Neifeld, and B Vasic. Bethe free energy approach to LDPC decoding on memory channels. cs/0904.0747 preprint, 2009.

[9] F Barahona. On the computational complexity of Ising spin glass models. *J. Phys. A*, 15:3241–3253, 1982.

[10] U Bastolla, H Frauenkron, and P Grassberger. Phase diagram of random heteropolymers. *Journal of Molecular Liquids*, 84:111–129, 2000.

[11] R J Baxter. *Exactly solved models in statistical mechanics*. ACADEMIC PRESS INC., 1982.

[12] D V Berkov. Local-field distribution in systems with dipolar interparticle interaction. *Phys. Rev. B*, 53:731, 1996.

[13] E R Berlekamp, R J McEliece, and H C A van Tilborg. On the inherent intractability of certain coding problems. *IEEE Trans. Inform. Theory*, 24:384, 1978.

[14] C Berrou, A Glavieux, and R Thitimajshima. Near Shannon limit error-correcting coding and decoding: Turbo codes. *Proc. IEEE Int. Conf. Commun.*, page 1064, 1993.

[15] K Binder and D W Heermann. *Monte Carlo Simulations in Statistical Physics*. Springer, 2002.

[16] D Bollé and R Heylen. Small-world hypergraphs on a bond-disordered Bethe lattice. *Phys. Rev. E*, 77:046104, 2008.

[17] D Bollé, R Heylen, and N S Skantzos. Thermodynamics of spin systems on small-world hyperghraphs. *Phys Rev E*, 74:056111, 2006.

[18] B Bollobás. *Random graphs*. Cambridge University Press, 2001.

[19] Jonathan M Borwein, D M Bradley, and Richard E Crandall. Computational strategies for the Riemann zeta function. *J. Comput. Appl. Math.*, 121:247–296.

[20] A Braunstein, A Pagnani, M Weigt, and R Zecchina. Inference algorithms for gene networks: a statisical mechanics analysis. *J. Stat. Mech.: Theory and Exp.*, 08:P12001, 2008.

[21] A J Bray and M A Moore. *J. Phys. C.: Solid St. Phys.*, 13:L4 69–76, 1980.

[22] Z Burda, J Jurkiewicz, M A Nowak, G Papp, and I Zahed. Free random lévy variables and financial probabilities. *Physica A*, 229:181–187, 2001.

[23] A Burr. Turbo-codes: the ultimate error control codes? *Electronics and Communication Engineering Journal*, pages 155–165, 2001.

[24] H B Callen. *Thermodynamics and introduction to thermostatics*. John Wiley and Sons, 1985.

[25] I Pérez Castillo, B Wemmenhove, J P L Hatchett, A C C Coolen, N S Skantzos, and T Nikoletopoulos. Analytic solution of attractor neural networks on scale-free graphs. *J. Phys. A: Math. Gen.*, 37:8789, 2004.

[26] P Cheeseman, B Kanefsky, and W M Taylor. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence*. IJCAI-91, 1999.

[27] S Ciliberti, T S Grigera, V Martín-Mayor, G Parisi, and P Verrocchio. Anderson localization in euclidean random matrices. *Phys. Rev. B*, 71:153104, 2005.

[28] P Cizeau and J P Bouchaud. Mean field theory of dilute spin-glasses with power-law interactions. *J. Phys. A: Math. Gen.*, 26:L187, 1993.

[29] P Cizeau and J P Bouchaud. Theory of Lévy matrices. *Phys. Rev. E*, 50:1810–1822, 1994.

[30] P Cizeau and J P Bouchaud. Theory of lévy matrices. *Phys. Rev. E*, 50:1810, 1994.

[31] L Correale, M Leone, A Pagnani, M Weigt, and R Zecchina. *J. Stat. Mech.: Theory and Exp.*, 03:P03002, 2006.

[32] T M Cover and J A Thomas, editors. *Elements of Information Theory.* John Wiley and Sons, Inc., Hoboken, New Jersey, 2006.

[33] F H C Crick. On protein synthesis. *Symp. Soc. Exp. Biol. XII*, pages 139–163, 1958.

[34] F H C Crick. Central dogma of molecular biology. *Nature*, 227:561–563, 1970.

[35] A Crisanti and H Sompolinsky. Dynamics of spin systems with randomly asymmetric bonds: Ising spins and Glauber dynamics. *Phys. Rev. A*, 12:4865–4874, 1988.

[36] J R L de Almeida and D J Thouless. Stability of the Sherrington-Kirkpatrick solution of a spin-glass model. *J. Phys. A: Math. Gen.*, 11:129, 1978.

[37] V M de Oliveira and J F Fontanari. Replica analysis of the p-spin interaction Ising spin-glass model. *J. Phys. A.: Math. Gen.*, 32:2285–2296, 1999.

[38] B Derrida. Random-energy model: An exactly solvable model of disordered systems. *Phys. Rev. B*, 24:2613–2626, 1981.

[39] B Derrida. Dynamical phase transition in non-symmetric spin glasses. *J Phys A: Math Gen*, 20:L721–L725, 1987.

[40] B Derrida, E Gardner, and A Zippelius. An exactly solvable asymmetric neural network model. *Europhysics Letters*, 4:167–173, 1987.

[41] S N Dorogovtsev and J F F Mendes. *Evolution of Networks: from biological networks to the Internet and WWW.* Oxford University Press, 2003.

[42] S N Dorogovtsev, J F F Mendes, and A N Samukhin. Structure of growing networks: Exact solution of the Barabási-Albert's model. *Phys. Rev. Lett.*, 85:4633, 2000.

[43] A W Eckford. *Low-Density Parity-Check Codes for Gilbert-Elliott and Markov-Modulated Channels*. PhD thesis, University of Toronto, 2004.

[44] A W Eckford, F R Kschischang, and S Pasupathy. Analysis of low-density parity-check codes for the Gilbert-Elliott channel. *IEEE Trans. Inform. Theory*, 51:3872–3889, 2005.

[45] T P Eggarter. Cayley trees, the ising problems, and the thermodynamic limit. *Phys. Rev. B*, 9:2989–2992, 1974.

[46] V M Eguíluz, D R Chialvo, G A Cecchi, M Baliki, and A Vania Apkarian. Scale-free brain functional networks. *Phys. Rev. Lett.*, 94:018102, 2005.

[47] H Eissfeller and M Opper. New method for studying the dynamics of disordered spin systems without finite-size effects. *Phys. Rev. Lett.*, 68:2094, 1992.

[48] E O Elliott. Estimates of error rates for codes on burst-noise channels. *Bell Syst. Tech. J.*, 42:1977–1997, 1963.

[49] P Erdős and A Rényi. On random graphs I. *Publc. Math. (Debrecen)*, 6:290, 1959.

[50] P Erdős and A Rényi. On the evolution of random graphs. *Magayr Tud. Akad. Mat. Kutato Int. Kozl.*, 5:17–61, 1960.

[51] L Giot et al. A protein interaction map of drosophila melanogaster. *Science*, 302:(5651):1727–36, 2003.

[52] P Uetz et al. A comprehensive analysis of protein-protein interactions in saccharomyces cerevisiae. *Nature*, 403:623–627, 2000.

[53] T Ito et al. Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc. Nat. Acad. Sci.*, 98:4569–1147, 2001.

[54] K H Fischer and J A Hertz. *Spin Glasses*, volume 1 of *Cambridge Studies in Magnetism*. Cambridge University Press, 1991.

[55] S Franz and M Leone. Replica bounds for optimization problems and diluted spin systems. *J. Stat. Phys.*, 111(3/4):535–564, 2003.

[56] S Franz, M Leone, A Montanari, and F Ricci-Tersenghi. The dynamic phase transition for decoding algorithms. *Phys. Rev. E*, 66:046120, 2002.

[57] S Franz, M Leone, and F L Toninelli. Replica bounds for diluted non-poissonian spin systems. *J. Phys. A: Math. Gen.*, 36:10967–10985, 2003.

[58] S Franz, M Mézard, F Ricci-Tersenghi, M Weigt, and R Zecchina. A ferromagnet with a glass transition. *Europhys. Lett.*, 55(4):465, 2001.

[59] C Furtlehner, J-M Lasgouttes, and A de La Fortelle. A belief propagation approach to traffic prediction using probe vehicles. *In Intelligent Transportation Systems Conference*, pages 1022–1027.

[60] R G Gallager. Low density parity check codes. *IRE Trans. Info. Theory*, 8:21, 1962.

[61] R G Gallager. *Low density parity-check codes*. MIT Press, Cambridge, 1963.

[62] J Garcia-Frias and J D Villasenor. Turbo decoding of Gilbert-Elliot channels,. *IEEE Trans. Commun.*, 50:357–363, 2002.

[63] E Gardner. Spin glasses with p-spin interaction. *Nucl. Phys. B*, 257:747–765, 1985.

[64] H O Georgii. *Gibbs measures and phase transitions*, volume 9 of *De Gruyter Studies in Mathematics*. de Gruyter Berlin, 1988.

[65] E N Gilbert. Capacity of a burst-noise channel. *Bell Syst. Tech. J.*, 39:1253–1265, 1960.

[66] P Gillin and D Sherrington. Comment on 'replica analysis of the p-spin interaction Ising spin-glass model'. *J. Phys. A.: Math. Gen.*, 34:1219–1222, 2001.

[67] A J Goldsmith and P P Varaiya. Capacity, mutual information, and coding for finite-state Markov channels. *IEEE Trans. Inform. Theory.*, 42:868–886, 1996.

[68] W Greiner, L Niese, and H Stöcker. *Thermodynamics and Statistical Mechanics*. Springer-Verlag, 1995.

[69] N Guelzim, S Bottani, P Bourgine, and F Kepes. Topological and causal structure of the yeast transcriptional regulatory network. *Nat. Genet.*, 31:60, 2002.

[70] F Guerra. Broken replica symmetry bounds in the mean-field spin glass model. *Commun math phys*, 233:1, 2003.

[71] F. Guerra and F L Toninelli. The thermodynamic limit in mean field spin glass models. *Commun. Math. Phys.*, 230:71, 2002.

[72] A K Hartmann and M Weigt. *Phase Transitions in Combinatorial Optimization Problems: Basics, Algorithms and Statistical Mechanics*, chapter 3 Introduction to graphs. Wiley-VCH, 2005.

[73] J P L Hatchett, I Pérez Castillo, A C C Coolen, and N S Skantzos. Dynamical replica analysis of disordered Ising spin systems on finitely connected random graph. *Phys. Rev. Lett.*, 95:117204, 2005.

[74] J P L Hatchett, B Wemmenhove, I Pérez Castillo, T Nikoletopoulos, N S Skantzos, and A C C Coolen. Parallel dynamics of disordered Ising spin systems on finitely connected random graphs. *J Phys A: Math Gen*, 37:6201, 2004.

[75] J Hertz, A Krogh, and R G Palmer. *Introduction to the Theory of Neural Computation*. Addison-Wesley.

[76] R A Horn and C R Johnson. *Matrix analysis*. Cambridge University Press, 1990.

[77] Y Iba. The Nishimori line and Bayesian statistics. *J. Phys. A: Math. Gen.*, 32:3875, 1999.

[78] K Janzen, A Engle, and M Mézard. The lévy spin glass transition. *Europhys. Lett.*, 89:67002, 2010.

[79] K Janzen, A K Hartmann, and A Engel. Replica theory for lévy spin glasses. *Journal of Statistical Mechanics: Theory and Experiment*, 26:04006, 2008.

[80] H Jeong, B Tombor, R Albert, Z N Oltvai, and A L Barabási. The large-scale organization of metabolic networks. *Nature*, 407:651–654, 2000.

[81] P Kaluza, A Kölzsch, M T Gastner, and B Blasius. The complex network of global cargo ship movements. arxiv:1001.2172.

[82] E F Keller. Revisiting scale-free networks. *BioEssays*, 27:1060–1068, 2005.

[83] S Kirkpatrick and D Sherrington. Infinite-ranged models of spin-glasses. *Phys. Rev. B*, 17:4384, 1978.

[84] T R Kirkpatrick and D Thirumalai. Dynamics of the structural glass transition and the p-spin-interaction spin-glass model. *Phys. Rev. Lett.*, 58:2091, 1987.

[85] T R Kirkpatrick and P Wolynes. Stable and metastable states in mean-field Potts and structural glasses. *Phys. Rev. B*, 36:8552, 1987.

[86] M W Klein. Temperature-dependent internal field distribution and magnetic susceptibility of a dilute Ising spin system. *Phys. Rev.*, 173:552, 1968.

[87] F R Kschischang, B J Frey, and H Loeliger. Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory*, 47:498–519, 2001.

[88] R Kühn. Spectra of sparse random matrices. *J. Phys. A: Math. Theory*, 41:295002, 2008.

[89] C Kwon and D J Thouless. Spin glass with two replicas on a Bethe lattice. *Phys. Rev. B*, 10:8379, 1991.

[90] M Leone and A Pagnani. Predicting protein functions with message passing algorithms. *Bioinformatics*, 21:239–247, 2005.

[91] M Leone, A Vazquez, A Vespignani, and R Zecchina. Ferromagnetic ordering in graphs with arbitrary degree distribution. *Eur. Phys. J. B*, 28:191, 2002.

[92] M Luby, M Mitzenmacher, M A Shokrollahi, and D Spielman. Analysis of low density codes and improved designs using irregular graphs. In *Proc. 30th Annu. ACM Symp. Theory of computing*, pages 249–258, 1998.

[93] M G Luby, M Mitzenmacher, M A Shokrollahi, and D A Spielman. Efficient erasure correcting codes. *IEEE Trans. Inform. Theory*, 47:569–584, 2001.

[94] D J C Mackay. Good error-correcting codes based on very sparse matrices. *IEEE Trans. IT*, 45:399, 1999.

[95] D J C MacKay. Errata for "good error-correcting codes based on very sparse matrices". *IEEE Transactions on Information Theory*, 47(5):2101, 2001.

[96] D J C Mackay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, The Pitt Building, Trumpington Street, Cambridge, UK, 2003.

[97] D J C MacKay and R M Neal. Near Shannon limit performance of low density parity check codes. *IEE Electronic Letters*, 32:1645, 1996.

[98] O C Martin, M Mézard, and O Rivoire. Frozen glass phase in the multi-index matching problem. *Phys. Rev. Lett.*, 93:217205, 2004.

[99] P C Martin, E Siggia, and H Rose. Statistical dynamics of classical systems. *Phys. Rev. A*, 8:423, 1973.

[100] M L Mehta. *Random matrices*. Elsevier Inc., 2004.

[101] F L Metz, I Neri, and D Bollé. On the localization transition in random matrices. http://arxiv.org/abs/1005.3712, 2010.

[102] M Mézard and A Montanari. *Information, Physics and Computation*. Oxford Graduate Texts, 2009.

[103] M Mézard and G Parisi. Mean-field theory of randomly frustrated systems with finite connectivity. *Europhys. Lett.*, 3:1067, 1987.

[104] M Mézard and G Parisi. The Bethe lattice spin glass revisited. *Eur. Phys. J. B*, 20:217, 2001.

[105] M Mézard and G Parisi. The cavity method at zero temperature. *J. Stat. Phys*, 1:11, 2003.

[106] M Mézard, G Parisi, N Sourlas, G Toulouse, and M A Virasoro. Nature of the spin-glass phase. *Phys. Rev. Lett.*, 52:1156, 1984.

[107] M Mézard, G Parisi, and M A Virasoro. *Spin Glass Theory and Beyond*, volume 9 of *World Scientific Lecture Notes in Physics*. World Scientific Pub Co Inc., 1987.

[108] G Migliorini and D Saad. Finite-connectivity spin-glass phase diagrams and low-density parity check codes. *Phys. Rev. E*, 73:026122, 2006.

[109] R Milo, S S Shen-Orr, S Itzkovitz, N Kashtan, D Chklovskii, and U Alon. Network motifs: Simple building blocks of complex networks. *Science*, 824:25, 2002.

[110] K Mimura and A C C Coolen. Generating functional analysi of LDGM channel coding with many short loops. to be published in Proceedings of the 2009 IEEE International Symposium on Information Theory (ISIT2009, Seoul, Korea), 509-513.

[111] K Mimura and A C C Coolen. Parallel dynamics of disordered Ising spin systems on finitely connected random graphs with arbitrary degree distributions. *Journal of Physics A: Mathematical and Theoretical*, 42:415001, 2009.

[112] M Molloy and B Reed. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, 6:161–179, 1995.

[113] R Monasson. Optimization problems and replica symmetry breaking in the finite connectivity spin glasses. *J. Phys. A.: Math. Gen.*, 31:513, 1998.

[114] A Montanari. The glassy phase of gallager codes. *Eur. Phys. J. B*, 23:121, 2001.

[115] A Montanari. Tight bounds for LDPC and LDGM codes under MAP decoding. *IEEE Trans. on Inf. Theory*, 51:3221, 2005.

[116] A Montanari, L F Cugliandolo, and G Semerjian. On the stochastic dynamics of disordered spin models. *J. Stat. Phys.*, 115:493–530, 2006.

[117] A Montanari and F Ricci-Tersenghi. On the nature of the low-temperature phase in discontinous mean-field spin glasses. *Eur. Phys. J. B*, 33:339–346, 2003.

[118] A Montanari and T Rizzo. How to compute loop corrections to the Bethe approximation. *J. Stat. Mech.: Theory Exp.*, page P10011, 2005.

[119] A Montanari and G Semerjian. On the dynamics of the glass transition on Bethe lattices. *J. Stat. Phys.*, 124:102, 2006.

[120] A Montanari and N Sourlas. The statistical mechanics of turbo codes. *Eur. Phys. J. B*, 18:107, 2000.

[121] T Mora and O Rivoire. Statistical mechanics of error exponents for error-correcting codes. *Phys. Rev. E*, 74:056110, 2006.

[122] A Mozeika and A C C Coolen. Dynamical replica analysis of processes on finitely connected random graphs i: vertex covering. *J. Phys. A: Math. Theor.*, 41:115003.

[123] E Müller-Hartmann and J Zittartz. New type of phase transition. *Phys. Rev. Lett.*, 33:893–897, 1974.

[124] T Murayama, Y Kabashima, D Saad, and R Vicente. Statistical physics of regular low-density parity-check error-correcting codes. *Phys. Rev. E*, 62:1577, 2000.

[125] M Mushkin and I Bar-David. Capacity and coding for the Gilbert-Elliot channels. *IEEE Trans. Inform. Theory.*, pages 1277–1290, 1989.

[126] I Neri and D Bollé. The cavity approach to parallel dynamics of Ising spins on a graph. *J. Stat. Mech.: Theory Exp.*, page P08009, 2009.

[127] I Neri, F L Metz, and D Bollé. The phase diagram of Lévy spin glasses. *J. Stat. Mech.: Theory and Exp.*, 10:P01010, 2010.

[128] I Neri and N S Skantzos. Statistical mechanics of LDPC codes on channels with memory. http://arxiv.org/abs/0907.0775, 2009.

[129] I Neri, N S Skantzos, and D Bollé. Gallager error-correcting codes for binary asymmetric channels. *J. Stat. Mech.*, page P10018, 2008.

[130] M E J Newman. The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. USA*, 98:404, 2001.

[131] T Nikoletopoulos and A C C Coolen. Diagonalization of replicated transfer matrices for disordered Ising spin systems. *J Phys A: Math Gen*, 37:8433, 2004.

[132] T Nikoletopoulos, A C C Coolen, I Perez-Castillo, N S Skantzos, J P L Hatchett, and B Wemmenhove. Replicated transfer matrix analysis of Ising spin models on 'small world' lattices'. *J Phys A: Math Gen*, 37:6455, 2004.

[133] H Nishimori. Optimum decoding temperature for error-correcting codes. *J. Phys. Soc. Japan*, 62:2973, 1993.

[134] H Nishimori. *Statistical physics of spin glasses and information processing*. Oxford Science publication, 2001.

[135] J P Nolan. *Stable Distributions: Models for Heavy-Tailed Data*. Birkhauser, 2007.

[136] A Pagnani, F Tria, and M Weigt. Classification and sparse-signature extraction from gene-expression data. *J. Stat. Mech.: Theory and Exp.*, 09:P05001, 2009.

[137] G Parisi. Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.*, 43:1754, 1979.

[138] G Parisi. Order parameter for spin-glasses. *Phys. Rev. Lett.*, 50:1946, 1983.

[139] G Parisi and F Tria. Spin glasses on Bethe lattices for large coordination number. *Eur. Phys. J. B*, 30:533–541, 2002.

[140] W Paul and J Baschnagel. *Stochastic Processes*. Springer-Verlag Berlin Heidelberg, 1999.

[141] M Potters, J P Bouchaud, and L Laloux. Financial applications of RMT : old laces and new pieces. *Acta Phys. Pol. B*, 36:2767, 2005.

[142] W H Press, S A Teukolsky, W T Vetterling, and B P Flannery. *Numerical Recipes in C*. Cambridge university press, 1994.

[143] A L Barabasi R Albert, H Jeong. The diameter of the world wide web. *Nature*, 401:130–131, 1999.

[144] J R Raymond and D Saad. Composite systems of dilute and dense couplings. *J. Phys. A: Math. Theor.*, 41:324014, 2008.

[145] J R Raymond and D Saad. Equilibrium properties of disordered spin models with two-scale interactions. *Phys. Rev. E*, 80:031138, 2009.

[146] T J Richardson, M A Shokrollahi, and R L Urbanke. Design of capacity-approaching irregular low-density parity-check codes. *IEEE Trans. Inform. Theory*, 47:619 – 637, 2001.

[147] T J Richardson and R L Urbanke. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Trans. Inform. Theory*, 47:599–618, 2001.

[148] T J Richardson and R L Urbanke. Efficient encoding of low-density parity-check codes. *IEEE Trans. on Inf. Theory*, 47:638–656, 2009.

[149] O Rivoire. *Phases Vitreuses, Optimisation et Grandes Déviations*. PhD thesis, Université Paris-Sud, 2005.

[150] T Rogers, I Pérez Castillo, R Kühn, and K Takeda. Cavity approach to the spectral density of sparse symmetric random matrices. *Phys. Rev. E.*, 78:031116, 2008.

[151] M A Ruderman and C Kittel. Indirect exchange coupling of nuclear magnetic moments by conduction electrons. *Phys. Rev.*, 96:99, 1954.

[152] P Sadeghi, R Kennedy, P Rapajic, and R Shams. Finite-state Markov modeling of fading channels - a survey of principles and applications. *IEEE Signal Processing Mag.*, 25:57–80, 2008.

[153] N Schwartz, R Cohen, D ben Avraham, A L Barabási, and S Havlin. Percolation in directed scale-free networks. *Phys. Rev. E*, page 015104, 2002.

[154] B Selman and S Kirkpatrick. *Artificial Intelligence*, 81:273–295, 1996.

[155] B Selman, D G Mitchell, and H J Levesque. *Artificial Intelligence*, 81:17–29, 1996.

[156] G Semerjian and L F Cugliandolo. Sparse random matrices: the eigenvalue spectrum revisited. *J. Phys. A: Math. Gen.*, 35:4837, 2002.

[157] G Semerjian and L F Cugliandolo. Dynamics of dilute disordered models: A solvable case. *Europhys. Lett.*, 61:247, 2003.

[158] E I Shakhnovich and A M Gutin. Formation of unique structure in polypeptide-chains theoritical investigation with the aid of a replica approach. *Biophys. Chem.*, 34:187–199, 1989.

[159] C E Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379, 623, 1948.

[160] S S Shen-Orr, R Milo, Sh Mangan, and U Alon. Network motifs in the transcriptional regulation network of Escherichia coli. *Nat. Genet.*, 31:64, 2002.

[161] D Sherrington and S Kirkpatrick. Solvable model of a spin glass. *Phys Rev Lett*, 35:1792, 1975.

[162] N S Skantzos, J van Mourik, D Saad, and Y Kabashima. Average and reliability error exponents in low-density parity-check codes. *J. Phys. A: Math. Gen.*, 36:11131, 2003.

[163] N Sourlas. Spin-glass models as error-correcting codes. *Nature*, 339:693, 1989.

[164] M Talagrand. The Parisi formula. *Ann. Math.*, 163:221–263, 2006.

[165] T Tanaka and D Saad. Typical performance of regular low-density parity-check codes over general symmetric channels. *J. Phys. A: Math. Gen.*, 36:11143, 2003.

[166] R M Tanner. A recursive approach to low complexity codes. *IEEE Trans. on Inf. Theory*, 27:533–547, 1981.

[167] S Tatikonda and M Jordan. Loopy belief propagation and gibbs measures. In Darwiche A and Friedman N, editors, *Uncertainty in artificial intelligence: Proceedings of the Eighteenth Conference (UAI-2002)*, pages 493–500. San Francisco: Morgan Kaufmann, 2002.

[168] H Touchette. The large deviation approach to statistical mechanics. *Phys. Rep.*, 8478:1–69, 2009.

[169] L Viana and A J Bray. Phase diagrams for dilute spin glasses. *J. Phys. C: Solid State Phys*, 18:3037, 1985.

[170] R Vicente, D Saad, and Y Kabashima. Finite-connectivity systems as error-correcting codes. *Phys. Rev. E*, 60:5352, 1999.

[171] R Vicente, D Saad, and Y Kabashima. Statistical physics of irregular low-density parity-check nodes. *J. Phys. A: Math. Gen.*, 33:6527, 2000.

[172] C Wang, S R Kulkarni, and H V Poor. Density evolution for asymmetric memoryless channels. *IEEE Trans. on Inf. Theory*, 51:4216, 2005.

[173] H S Wang. On verifying the first-order Markovian assumption for a Rayleigh fading channel model. *IEEE Trans. Veh. Technol.*, 44:163–171, 1995.

[174] D J Watts and S H Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998.

[175] J White, E Southgate, J Thomson, and S Brenner. The structure of the nervous system of the nematode C. elegans. *Phil Trans R Soc Lond B*, 314:1–340, 1986.

[176] J S Yedidia, W T Freeman, and Y Weiss. *Understanding Belief Propagation and its Generalizations, Exploring Artificial Intelligence in the New Millennium*, chapter 8, pages 239–271. Science and Technology Books, 2003.

[177] L Zdeborová. *Statistical Physics of Hard Optimization Problems*. PhD thesis, Univerzita Karlova v Praze, 2008. http://arxiv.org/abs/0806.4112.

[178] L Zdeborová and F Krzakala. Phase transitions in the coloring of random graphs. *Phys. Rev. E*, 76:031131, 2007.

Curriculum vitae

Personal data

May 25 1983, Deinze
Male
Belgian

Institute of Theoretical Physics, K.U.Leuven
Celestijnenlaan 200D
B-3001 Leuven
+32-16-3 27236 (office)
izaakneri@gmail.com

Education

Ph. D. in theoretical physics
K. U. Leuven, February 2006 - May 2010
Statistical mechanics of spin models on graphs
Supervisor: Prof. Dr. D. Bollé
January-February 2008: visit at King's college London

Teaching assistant
statistical mechanics (2006-2010), biophysics (2010), physics for architects (2009), artificial neural networks (2008)

M. Sc. in physics
UGent, September 2003- July 2005
Master thesis: High Tc superconductivity and the Hubbard model

Bachelor in physics
UGent, September 2001- July 2003

Published and submitted articles

1. F L Metz, I Neri and D Bollé
On the localization transition in random matrices
arXiv:1005.3712

2. The phase diagram of Lévy spin glasses
   I Neri , F L Metz and D Bollé
   Journal of Statistical Mechanics: Theory and Experiment (2010) P01010

3. I Neri and N S Skantzos
   Statistical mechanics of LDPC codes on channels with memory
   arXiv:0907.0775

4. I Neri and D Bollé
   The cavity approach to parallel dynamics of Ising Spins on a graph
   Journal of Statistical Mechanics: Theory and Experiment (2009) P08009

5. I Neri, N S Skantzos and D Bollé
   Gallager error-correcting codes for binary asymmetric channels
   Journal of Statistical Mechanics: Theory and Experiment (2008) P10018

Conferences

- STATPHYS, 2007 (Genova)
- Conference of the Middle European Cooperation in Statistical Physics, 2008 (Puchberg-Wels)
  Poster: Gallager codes on asymmetric codes
- European Conference on Complex Systems, 2008 (Jerusalem)
- Kavli Institute for Theoretical Physics China, 2010 (Beijing)
  Statistical Physics and Computer Science
  Talk: Steady states of spin models on graphs

Arenberg Doctoral School of Science, Engineering & Technology
Faculty of Science
Department of Physics and Astronomy
Physics of Disordered Systems
Celestijnenlaan 200D
B-3001 Leuven

KATHOLIEKE UNIVERSITEIT LEUVEN

ASSOCIATIE K.U.LEUVEN